

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

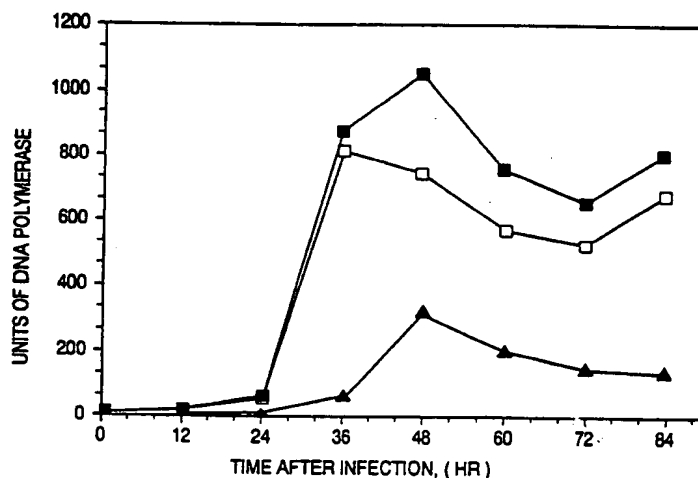


INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C07H 21/04, C12P 21/02 C12N 15/86, C12Q 1/48, 1/70	A1	(11) International Publication Number: WO 93/10139 (43) International Publication Date: 27 May 1993 (27.05.93)
--	----	--

(21) International Application Number: PCT/US92/09457

(22) International Filing Date: 12 November 1992 (12.11.92)

(30) Priority data:
792,600 15 November 1991 (15.11.91) US(71) Applicant: BOARD OF TRUSTEES OF THE LELAND
STANFORD JR. UNIVERSITY [US/US]; Stanford
University, Standord, CA 94305-6225 (US).(72) Inventors: COPELAND, William, C. ; 3740 Grove Avenue,
Palo Alto, CA 94303 (US). WANG, Teresa, S., F. ; 701
Tennyson Avenue, Palo Alto, CA 94303 (US).(74) Agents: CARROLL, Peter, G. et al.; Haverstock, Medlen
& Carroll, 220 Montgomery Street, Suite 2200, San Fran-
cisco, CA 94104 (US).(81) Designated States: AU, CA, JP, European patent (AT, BE,
CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL,
SE).Published
With international search report.(54) Title: COMPOSITIONS AND METHODS FOR TEMPLATE-DEPENDENT ENZYMATIC SYNTHESIS OF NUC-
LEIC ACID

(57) Abstract

The human DNA polymerase α catalytic polypeptide has been functionally over-expressed by a recombinant baculovirus in insect cells at ~ 1000 fold higher levels than that found in cultured normal human cells.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	LI	Liechtenstein	SK	Slovak Republic
CI	Côte d'Ivoire	LK	Sri Lanka	SN	Senegal
CM	Cameroon	LU	Luxembourg	SU	Soviet Union
CS	Czechoslovakia	MC	Monaco	TD	Chad
CZ	Czech Republic	MG	Madagascar	TG	Togo
DE	Germany	ML	Mali	UA	Ukraine
DK	Denmark	MN	Mongolia	US	United States of America
ES	Spain			VN	Viet Nam
FI	Finland				

-1-

DescriptionCOMPOSITIONS AND METHODS FOR TEMPLATE-DEPENDENT
ENZYMATIC SYNTHESIS OF NUCLEIC ACIDField of the Invention

5 The present invention relates to compositions
and methods for template-dependent enzymatic
synthesis of nucleic acid, and more specifically,
replication of nucleic acid by human polymerase α .
The present invention is particularly useful for
10 screening chemotherapeutics for potential
mutagenicity and carcinogenicity.

Background

 The term "chemotherapy" simply means the
treatment of disease with chemical substances. The
15 father of chemotherapy, Paul Ehrlich, imagined the
perfect chemotherapeutic as a "magic bullet"; such a
compound would kill an invading organism without
harming the host. This target specificity is sought
in all types of chemotherapeutics, including
20 antimicrobial and anticancer agents.

 Unquestionably, the greatest success with
antimicrobials in terms of specificity has been with
antibiotics. The antibiotic penicillin is widely
known for its ability to block the synthesis of the
25 cell wall for particular bacteria without interfering
with the biochemistry of mammalian cells. What is
not widely known is that penicillin is the exception
rather than the rule; only a fraction of the
thousands of identified antimicrobial drugs are non-
toxic to humans.
30

 Efforts to treat viral infection have been
largely ineffective for precisely this reason. While
a virus is essentially nothing more than nucleic acid

-2-

surrounded by a lipid-protein envelope, a virus
invades a host cell and uses the host cell's
machinery to replicate itself. The latter
characteristic makes it especially difficult to find
5 drugs which block viral replication and yet leave
intact the ability of the host cell to replicate.

Specificity has also been the major problem with
anticancer agents. In the case of anticancer agents,
the drug needs to distinguish between host cells that
10 are cancerous and host cells that are not cancerous.
The vast bulk of anticancer drugs are indiscriminate
at this level. For this reason, only a few types of
cancer are appropriate for chemotherapy. Surgery and
radiation continue to be the favored types of cancer
15 treatment.

Drug Screening

While there has been little success with viral
infection and cancer, there is continued hope that
drugs can be found or designed with the requisite
20 specificity for the treatment of human afflictions.
However, even if compounds can be found that do not
have immediate toxicity, exhaustive screening is
necessary to ensure that the selected compounds are
neither carcinogens nor mutagens.

25 A mutation is a change in the sequence, number
or nature of nucleotide bases in DNA. A certain
amount of mutation is normal (and perhaps even
necessary) in all organisms. A mutagen is a compound
that increases the normal frequency of mutation.

30 One source of mutation is caused by direct
modification of a normal base by a mutagen so as to
alter its normal base pairing. Another type of
mutation is caused by the incorporation of analogs of
the normal nucleotide bases during DNA replication.

-3-

Still other mutations are caused by the incorporation of additional bases or the loss of bases during replication.

Importantly, not all mutagens will result in carcinogenicity. Nonetheless, all carcinogens are mutagens.

It has proven difficult to directly measure mutagenicity of compounds in higher organisms such as mammals. Mutations are rare and it takes great numbers of organisms before they are seen. Current approaches, therefore, utilize microorganisms such as bacteria.

The most widely used mutagen/carcinogen screening assay is the Ames test. The Ames test utilizes several unique strains of *Salmonella typhimurium* that are histidine-dependent for growth and that lack the usual DNA repair enzymes. The frequency of normal mutations that render the bacteria independent of histidine (i.e., the frequency of spontaneous revertants) is low. Thus, the test can evaluate the impact of a compound on this revertant frequency.

Since some substances are not mutagenic by themselves but are converted to a mutagen by metabolic activation, the compound to be tested is mixed with the bacteria on agar plates along with a liver extract. The liver extract is needed to mimic metabolic activation in an animal. Control plates have only the bacteria and the extract.

The mixtures are allowed to incubate. Growth of bacteria (if any) is checked by counting colonies. A positive Ames test is one where the number of colonies on the plates with mixtures containing the compound significantly exceeds the number on the corresponding control plates.

-4-

When known carcinogens are screened in this manner with the Ames test, approximately ninety percent are positive. When known noncarcinogens are similarly tested, approximately ninety percent are negative.

Drawbacks to the Bacterial Model

For many compounds, the Ames test is quite adequate. These compounds (e.g., pesticides, dyes, etc.) are those thought to cause mutations by direct modification of the chemistry of a normal base. It is believed that this nucleic acid modification chemistry will be the same in the bacteria as in mammalian cells. Thus, the change in the revertant frequency of the bacteria is predictive of mutagenicity in mammals.

The Ames test is, however, not definitive for all chemotherapeutics. Indeed, it may be particularly ill-suited to test nucleotide analogs designed as antiviral and anticancer agents. These agents are designed to be incorporated in the target cell nucleic acid during replication. Unfortunately, they may also be incorporated by normal host cells during normal replication and cause subsequent mutations.

In contrast to nucleic acid modification chemistry, incorporation of nucleotide analogs occurs via the replication machinery. It is known that the bacterial replication machinery is distinctly different from that of mammalian cells. Consequently, there is a concern that there will be a class of nucleotide analogs that will not be incorporated by bacteria but that will be incorporated by normal replicating mammalian cells. These compounds would

-5-

test negative by the Ames test and yet be mutagenic in mammals.

SUMMARY OF THE INVENTION

5 The present invention relates to compositions and methods for template-dependent enzymatic synthesis of nucleic acid, and more specifically, replication of nucleic acid by human polymerase α . The use of human DNA polymerase α is particularly appropriate for screening chemotherapeutics for potential mutagenicity and carcinogenicity. Unlike
10 current screening approaches, the screening approach of the present invention is predictive of the mutagenicity (if any) of nucleotide analogs in mammals.

15 The present invention contemplates the over-expression of recombinant human DNA polymerase α that is functional, and yet free of contaminating protein typically associated with human DNA polymerase α purified by traditional biochemical isolation
20 techniques. The expression of recombinant human DNA polymerase α of the present invention relies on the construction of a full-length cDNA. This full-length cDNA has been found to generate full-length translation products.

25 The present invention contemplates the use of recombinant human DNA polymerase α for the screening of chemotherapeutics for potential mutagenicity and carcinogenicity. In one embodiment, recombinant human DNA polymerase α is employed to test for
30 incorporation of analogs of the normal nucleotide bases during DNA replication.

 The present invention further contemplates the use of recombinant human DNA polymerase α to test for the binding of viral proteins. In one embodiment,

-6-

the present invention contemplates co-infection of cells with two expression vectors, one vector coding for the viral protein of interest and the other vector coding for human DNA polymerase α .

5 It is not intended that the present invention be limited by the expression system for recombinant human DNA polymerase α . The present invention contemplates all forms and sources of expression systems.

10 DESCRIPTION OF THE DRAWINGS

Figure 1 is the restriction map of the cDNA for human DNA polymerase α , showing overlapping cDNA clones (triangles designate the locations of the frame-shift mutations).

15 Figure 2 is a photograph of an autoradiograph, following in vitro translations used to confirm the correct sequence of the full-length cDNA for human DNA polymerase α .

20 Figure 3 sets forth the correct nucleotide and amino acid sequence of the human DNA polymerase α catalytic polypeptide (newly corrected sequences are boxed in).

Figure 4 schematically shows the pBlueBac/HDP α expression system.

25 Figure 5 shows the level of expression of human DNA polymerase α at intervals, post-infection. Figures 5A, 5C, 5D and 5E are photographs of Coomassie-stained gels following electrophoresis (SDS-PAGE). Figure 5F schematically shows the enzymatic activity by polymerase assays. Figure 5B is an immunoblot.

30

Figure 6 shows the analysis of the recombinant polymerase α of the present invention by immunoprecipitation and phosphorylation. Figure 6A

-7-

is a photograph of an SDS-PAGE gel stained with Coomassie blue following electrophoresis. Figure 6B is a photograph of an autoradiogram of lanes 3 and 4 of Figure 6A.

5 Figure 7 shows the protein profile during immunoaffinity purification of the single subunit recombinant human polymerase α of the present invention. Figure 7A is a photograph of a Coomassie stained gel following electrophoresis. Figure 7B is
10 a photograph of a silver stained gel following electrophoresis.

 Figure 8 shows the absence of a 3'-5' proofreading exonuclease activity in the single subunit recombinant polymerase α . Figure 8A is an
15 autoradiogram of the exonuclease assay using 0.1 pmole of the ^{32}P -5'-labeled mismatched 29mer annealed to M13mp18 single stranded DNA as substrate. Figure 8B is an autoradiogram of the exonuclease assay using 0.1 pmole of the ^{32}P -5'-labeled matched
20 24mer annealed to M13mp18 single stranded DNA as substrate.

 Figure 9 is an autoradiogram following electrophoresis showing the lack of exonuclease activity of the polymerase of the present invention
25 on a polynucleotide template.

 Figure 10 shows a comparative analysis of DNA synthetic processivity on primed M13 single-stranded DNA and oligo(dT) primed poly(dA).

 Figure 11 shows the use of the polymerase of the present invention as a reagent for chemotherapeutic
30 screening. Figure 11A is an autoradiogram following electrophoresis of "standing start" primer-templates used to test for the incorporation of either dTTP or AZTTP. Figure 11B is an autoradiogram following
35 electrophoresis of "running start" primer-templates

-8-

used to test for the incorporation of either dTTP or AZTTP.

Figure 12 shows an example of the interaction of recombinant human DNA polymerase α of the present invention with viral proteins. Figure 12A is a Coomassie Blue stained gel of cell lysates and anti-polymerase α immunoprecipitations. Figure 12B is a (Western) immunoblot.

DESCRIPTION OF THE INVENTION

The present invention relates to compositions and methods for template-dependent enzymatic synthesis of nucleic acid. More specifically, the present invention relates to replication of nucleic acid by human polymerase α .

It is believed that eukaryotic DNA replication requires at least two DNA polymerases, α and δ , for the lagging and leading strand synthesis, respectively. See B. Stillman, Ann. Rev. Cell Biol. 5:197 (1989). M.D. Challberg and T.J. Kelly, Ann. Rev. Biochem. 58:671 (1989). DNA polymerase α /primase complex is responsible for the synthesis of the nascent DNA fragment during initiation of DNA replication and for lagging strand DNA synthesis during elongation. Evidence from several laboratories suggests that interactions between DNA polymerase α and other replication proteins are highly stringent and species specific. See T. Tsurimoto *et al.*, Nature 346:534 (1990).

Studies of the structure and biological function of DNA polymerase α have been problematic due to its low abundance in cells and susceptibility to proteolysis during purification. Nonetheless, development of immunoaffinity and biochemical purification protocols in recent years has allowed

-9-

the demonstration that polymerase α activity purified from a wide phylogenetic range of species contains a similar set of constituent subunit components. The enzyme complex is made up of a cluster of large phosphopolypeptides of predominantly 165 to 180 kDa with catalytic function, a 70 kDa phosphoprotein of unknown function, and two polypeptides, 55 and 49 kDa, containing the primase activity. This four subunit component containing polymerase α and primase activities has been designated polymerase α /primase complex. T. Wang, Ann Rev. Biochem. 60:513 (1991). Peptide mapping of the p180 and p165 subunits indicate that they are derivatives of the same polypeptide. S.W. Wong et al., J. Biol. Chem. 261:7958 (1986).

The present disclosure describes the isolation of the correct, full-length cDNA of the catalytic polypeptide of human DNA polymerase α . The full length human cDNA has been constructed to be functionally over-expressed in a baculovirus transfer vector for expression in insect cells. However, it also is constructed to be functionally over-expressed in monkey COS7 cells and in yeast. Indeed, it is not intended that the present invention be limited by the expression system for recombinant human DNA polymerase α . The present invention contemplates all forms and sources of expression systems.

Importantly, the human DNA polymerase α catalytic polypeptide has been functionally over-expressed at >1000 fold higher levels than that found in cultured normal human cells. The recombinant polymerase α protein is translated from its natural translation start codon producing a protein of 180 kDa, identical in size to that isolated from cultured human cells. This recombinant polymerase α ,

-10-

immunopurified as a single polypeptide, is phosphorylated and reactive to a panel of monoclonal antibodies directed against the native polymerase α /primase complex and to polyclonal antisera against N- and C-terminal peptides of the polymerase α catalytic polypeptide. The single subunit recombinant polymerase α has no detectable 3'-5' exonuclease activity. The k_m for primer-template and dNTP, reactivity to inhibitors, thermosensitivity, and DNA synthetic processivity and fidelity of the recombinant polymerase α are identical to that observed with the four subunit polymerase α /primase complex immunopurified from cultured human cells.

The present invention contemplates using human polymerase α for template-dependent enzymatic synthesis of nucleic acid, and more specifically, for replication of nucleic acid. The present invention is useful for screening chemotherapeutics for potential mutagenicity and carcinogenicity. As noted above, one type of mutation is caused by the incorporation of analogs of the normal nucleotide bases during DNA replication. The present invention is particularly useful for screening chemotherapeutics that are analogs of the normal nucleotide bases.

The present invention further contemplates the use of recombinant human DNA polymerase α to test for the binding of viral proteins. In one embodiment, the present invention contemplates co-infection of cells with two expression vectors, one vector coding for the viral protein of interest and the other vector coding for human DNA polymerase α .

-11-

DETAILED DESCRIPTION OF THE INVENTION

The detailed description of the invention is divided into seven major sections: I) cDNA Construction, II) Protein Expression, III) Protein Purification and Characterization, IV) Template-Dependent Enzymatic Synthesis, V) Chemotherapeutic Screening, VI) Drug Design and VII) Viral Protein Binding.

I. cDNA CONSTRUCTION

Construction of cDNA was performed as generally described by Maniatis et al., Molecular Cloning, A Laboratory Manual, (Cold Spring Harbor Press, NY 1982). Essentially, this involved a) biochemical isolation of the polymerase, b) peptide sequencing, c) probe design, d) preparation of a cDNA library, e) screening of the library with the probes, f) isolation of the positive clone(s), and g) sequencing of the cDNA.

Biochemical Isolation. Biochemical isolation of human polymerase α is well-known. See e.g., F.J. Bollum, J. Biol. Chem. 235:2399 (1960). M. Mechali et al., J. Biol. Chem. 255:2114 (1980). L.S. Kaguni et al., Proc. Natl. Acad. Sci. USA 80:2221 (1983). L. Chang et al., J. Biol. Chem. 259:14679 (1984). P. Plevani et al., J. Biol. Chem. 260:7102 (1985). F. Grosse and G. Krauss, J. Biol. Chem. 260:1881 (1985). S.W. Wong et al., J. Biol. Chem. 261:7958 (1986). R. Lehman and L.S. Kaguni, J. Biol. Chem. 264:4265 (1989). However, progress in understanding the structure and properties of the native enzyme has been severely hindered by the low abundance of the enzyme, the apparent complexity and heterogeneity of the polymerase activity in impure fractions, and the

-12-

arduous purification schemes that have resulted in polymerase inactivation.

In this case, biochemical isolation was as generally described by T. Wang *et al.*, J. Biol. Chem. 259:1854 (1984) and S.W. Wong *et al.*, J. Biol. Chem. 261:7958 (1986). The KB cell line was used; this is a human epidermoid carcinoma cell line available from the American Type Culture Collection (ATCC) (Rockville, MD). The isolation of human DNA polymerase α catalytic polypeptides from this cell line proceeded as follows. DNA polymerase α antigen polypeptides from six, 18 liter cultures of human KB cells (3.5×10^5 cells/ml) were purified with a monoclonal IgG-Sepharose 4B column. The monoclonal, SJK287, was raised against a biochemically-purified, catalytically-active polymerase preparation. See S. Tanaka *et al.*, J. Biol. Chem. 257:8386 (1982). The polypeptides were suspended in 0.36 M Tris-HCl, pH 8.6, 3.3 mM EDTA, 8 M urea and then reduced for 3 h at 37°C under N_2 with 10 mM dithiothreitol. The reduced polypeptides were alkylated with 22 mM iodoacetic acid at 4°C for 1 h and dialyzed in 50 mM NH_4HCO_3 , 0.01% SDS. The dialyzed, reduced and alkylated DNA polymerase α protein was lyophilized, resuspended in 100 mM $NaPO_4$, pH 6.5, and 0.1% SDS and heated at 75°C for 10 min. These polypeptides were then purified by HPLC through two coupled gel permeation columns (TSK 3000, 7.5 x 300 mm) in 100 mM $NaPO_4$, pH 6.5, and 0.1% SDS at a flow rate of 0.5 ml/min. The absorbance of the eluate was monitored at 280 nm. Fractions containing the 180-140 kd DNA polymerase α catalytic polypeptides were dialyzed in 50mM NH_4HCO_3 containing 0.01% SDS and lyophilized.

-13-

Peptide sequencing. Peptide sequence analysis was performed as follows. Human DNA polymerase α catalytic polypeptides (500 pmol), isolated as described above, were resuspended in H_2O and ethanol-precipitated twice to remove excess SDS from the samples. The polypeptides were then resuspended in 0.1 M NH_4HCO_3 , 10 mM $CaCl_2$ and digested with 2 μg of TPCK treated trypsin at room temperature for 20 h. The trypsin digested peptides were first separated on an Aquapore RP300 (2.1 x 220 mm, Brownlee Lab) HPLC column equilibrated in 0.1% trifluoroacetic acid. A linear gradient from 0-60% acetonitrile was run over 45 min. at 0.2 ml/min. Absorbance at 220 nm was monitored by Spectraflow 755 Variable Wavelength detector. Selected peptides peaks were further purified by an RP300 (1 x 100 mm) column equilibrated in 50 mM ammonium acetate, pH 6.5. A linear gradient of 0-75% acetonitrile was run over 30 min at 0.08 ml/min. and absorbance monitored at 215 nm. Each of the separated peptides was subjected to automated Edman degradation performed on a model 470A gas phase sequencer with on-line PTH amino acid analysis (Model 120A) (Hunkapiller et al., 1983).

The amino acid sequences of seven peptides (hereinafter designated T9, T19, T23, T24, T25, T264, and T265) were determined as described in Table 1 of S.W. Wong et al., The EMBO Journal 7:37 (1988). In all, the sequences of 85 amino acids were established.

Probe design. Using the amino acid data, single long anti-sense oligonucleotide probes were designed according to R. Lathe, J. Mol. Biol. 183:1 (1985). The probes were synthesized on an Applied Biosystems model 380A oligonucleotide synthesizer at DNAX

-14-

Research Institute (the degenerate code is as follows: 3=C/G, 4=A/T, 5=A/G, 6=C/T, 7=A/C/G/T, 8=A/C/T, 9-A/C):

- 5 Polalpha #19A
Position: peptide T19
Sequence: GCTGCCTATGCTGGCGGCCTGGTGCTGGACCCAAG
- Polalpha #19B
Position: peptide T19
Sequence: CTTGGGGTCCAGCACCAGGCCGCCAGCATAGGCAGC
- 10 Polalpha #23A
Position: peptide T23
Sequence: CTTACCTCCAGCCAGGTGGGGCC
- Polalpha 25
Position: peptide T25
15 Sequence: TA6AT8TT6GA6GC7GA
- Polalpha 25
Position: peptide T25
Sequence: TACATCTTTGATGCTGAGACAGCCCTGGAGAAG
- 20 Polalpha 25A
Position: peptide T25
Sequence: CTTCTCCAGGGCTGTCTCAGCATCAAAGATGTA
- Polalpha 26A
Position: peptide T26
Sequence: GTAGAACACCTGCTGCAGCAGCTCATC
- 25 Polalpha #23AI
Position: peptide T23
Sequence: ITTIACITCIA5CCAIGTIGGICC
- Polalpha #24AI
Position: peptide T24
30 Sequence: ITTITCIGGIACITCIGGIATITCIAAIGCITAITT
- Polalpha #25AI
Position: peptide T25
Sequence: ITTITCIA5IGCIGTITCIGCITCIAAIATITA
- 35 Polalpha #26AI
Position: peptide T26
Sequence: ITTITCIGGIACITCIGGIATITCIAAIGCITAITT
- P264a
Position: peptide T264
Sequence: AC7GG7AA6TT6GT

-15-

P265

Position: peptide T265

Sequence: GATCTGCTGGGCCAGGTAGTACTGGG
TGTC AATGGTCAGGTTGGTCTG

5

T-26-4

Position: peptide T264

CCGGGACTGGTCAGACAGGATCTGGCCAATCACAAAGTTGCCTGT

These probes were used to screen cDNA libraries for positive clones.

10

Preparation of a cDNA library. Ninety μ g of poly(A)⁺ mRNA from early mid-log human KB cells was heated at 65°C for 1 min. and loaded onto a 5.3 ml sucrose gradient of 5-25% containing 100 mM NaCl, 10 mM Tris-HCl, pH 7.4, 1 mM EDTA and 0.1% SDS. Centrifugation was carried out at 52 300 g for 25 h at 5°C and fractionated into 20 fractions. mRNA samples of each fraction were precipitated by ethanol and resuspended into 5 μ l of H₂O. Ten percent of each fraction was used to estimate the size by reverse transcription, followed by analyzing the product on a 1% alkaline agarose gel. Fractions containing mRNA of >4 kb were used to construct a cDNA library in pCD vector as described. See H. Okayama and P. Berg, Mol. Cell. Biol. 2:161 (1982) and 3:280 (1983). An aliquot of this library containing 1×10^5 recombinants was used for the screening with oligonucleotide probes.

15

20

25

Screening of the library. The hybridization conditions used for screening were 6 x SSPE, 0.1% SDS 100 μ g/ml E. coli tRNA. Washing conditions were 2 x SSPE, 0.1% SDS. Temperature of hybridization and washing depended on the individual oligonucleotide probe used. Stringency of hybridization and washing of each individual oligonucleotide probe was based on

30

-16-

T_m (melting temperature) and T_w (washing temperature) values estimated at >85% probe-target homology. See R. Lathe, J. Mol. Biol. 183:1 (1985).

Isolation of the positive clones. Screening of
5 1×10^5 colonies of this size-selected library yielded a single distinct positive clone designated as pcD-KBpol α , which hybridizes with oligo-deoxynucleotide probes corresponding to peptides T264, T265 and T25 (see above).

Sequencing of the cDNA. Preliminary sequence
10 analysis of pcD-KBpol α indicated that it contained a 2893-bp cDNA insert with an open reading frame of 1865 bp terminated by a stop codon and followed by a 1028-bp non-coding region. See S.W. Wong et al., The
15 EMBO Journal 7:37 (1988). In this 1865-bp coding sequence there are four regions of deduced amino acid sequences that appeared perfectly homologous to the previously determined amino acid sequences, T264, T265, T25 and T9. The 3'-non-translated region
20 contains several in-frame stop codons, and the consensus polyadenylation signal AATAAA (N.J. Proudfoot and G.G. Brownlee, Nature 263:211 1976) 13 nucleotides upstream from the polyadenylation tail. This strongly suggested that pcD-KBpol α contained the
25 3'-end of the cDNA for human DNA polymerase α .

To extend this truncated cDNA clone the 5'-most restriction fragment of pcD-KBpol α , PstI/HindIII, was used to screen 2×10^6 phage of a human pre-B cell cDNA library (El library) constructed in λ gt10. See
30 Cleary et al., Cell 47:19 (1986). The very 5'-terminal restriction fragments of the newly extended cDNA clones were used to further screen the El library. Some of the clones were sequenced in both

-17-

directions as described (R.M.K. Dale et al., Plasmid 13:31 1985).

II. PROTEIN EXPRESSION

Initial attempts to functionally express the full length cDNA clone of the human DNA polymerase α catalytic subunit resulted in truncated translation products. Resequencing of the five overlapping cDNA clones in conjunction with in vitro translation analysis revealed two frame-shift mutations and two missense mutations in the two previously isolated cDNA clones, E1-19 and E1-12, that contain the 5' end of the cDNA sequence. Figure 1 is the restriction map of the cDNA for human DNA polymerase α , showing overlapping cDNA clones (triangles designate the locations of the frame-shift mutations).

The deletion frame-shift in E1-19 shifted translation by +1 at the nucleotide 1336 and caused termination of the protein after translation of nucleotides 1419-1421 (TGA). Figure 2 is a photograph of an autoradiograph, following in vitro translations used to confirm the correct sequence of the full-length cDNA for human DNA polymerase α . Translation products were labelled with [35S]-L-methionine and subjected to electrophoresis (SDS-PAGE, 10% gels) followed by autoradiography for four days. Positive and negative controls are in Lanes 1 and 2, respectively. Lane 4 shows the translation (in vitro translations) of the cDNA containing the E1-19 cDNA clone, producing an apparent polypeptide of 76,000 daltons in the rabbit reticulocyte lysate system (Promega Corp., Madison, Wisconsin). This region of the E1-19 cDNA clone was not sequenced previously (see S.W. Wong et al., The EMBO Journal 7:37 1988) and thus went undetected.

-18-

The insertion frame-shift in the E1-12 clone when spliced to the first half of the E1-19 clone shifted translation by -1 at the stretch of 6 A's causing termination in protein synthesis after nucleotides 1558-1560 (TAA). This gave an apparent translation product in vitro of 84,000 daltons (see lane 3 of Figure 2). Previously (see S.W. Wong et al., The EMBO Journal 7:37 1988) the outlined G was dropped from the sequence which shifted the predicted reading frame back in frame. In this area of the sequence only the E1-12 clone was sequenced so there was no comparison to the correct sequence found in the E1-19 clone.

To correct the frame-shifts it was easier to change the mutation in the E1-19 clone and use that clone with E1-14a to reconstruct the full length cDNA. The alternative would have been to splice the first third of the E1-19 clone with a small portion of E1-12 covering the E1-19 mutation, and then splicing the last part of E1-19 to cover the E1-12 mutation. Because of the close proximity of these mutations it was much easier to change the E1-19 mutation by site-directed mutagenesis using a custom designed oligo (hereinafter "BC2"): 5'-G AAC TAT GCA TTC GAG ATA CCT GA-3'. BC2 also contains an NsiI restriction endonuclease site for monitoring the presence of the mutation throughout the subcloning steps.

Briefly, the BC2 oligo was annealed to single stranded M13 phage uracil rich DNA containing the 1505bp PstI-PstI fragment of E1-19 and extended by T4 DNA polymerase and selected in JM101 E. coli. The correct mutation was confirmed by DNA sequencing and reconstructed into the E1-19 cDNA clone. The full length cDNA clone was reconstructed from the

-19-

corrected E1-19 clone and the E1-14a clone by splicing them together at the Sall site at nucleotide position 2004. Lanes 5 and 6 of Figure 2 show the products from translation of the full length cDNA with the corrected E1-19 cDNA clone. The same translations were performed in lanes 7 and 8 of Figure 2, but these were performed in the presence of caffeine to stimulate full length translation. Note the production of a band at 180 kDa (top arrow).

As a result of these findings and according to the sequence data of the panel of overlapping cDNA clones, the amino acid sequence of the human DNA polymerase α was corrected as follows: the previously reported amino acids KSTA from amino acid residue position 499 to 503 are changed to SPQL, and amino acid residue G at position 837 have been corrected to A by site-directed mutagenesis to give a continuous open reading frame of 1462 amino acids. Figure 3 sets forth the correct nucleotide and amino acid sequence of the human DNA polymerase α catalytic polypeptide (newly corrected sequences are boxed in).

In one embodiment, protein expression is carried out using a recombinant baculovirus expression vector, capable of expression in a host insect cell. Such systems are known to the art. For example, G.E. Smith and M.D. Summers, U.S. Patent Nos. 4,745,051 and 4,879,236, hereby incorporated by reference, describe a method wherein baculovirus DNA is cleaved to produce a DNA fragment comprising a polyhedrin gene, including a polyhedrin promoter. A recombinant shuttle vector is prepared by inserting the fragment into a cloning vehicle and thereafter inserting a selected gene into the modified cloning vehicle such that it is under the transcriptional control of the polyhedrin promoter. The recombinant shuttle vector

-20-

is contacted with a baculovirus DNA so as to effect recombination and incorporation of the selected gene into the baculovirus genome. The resultant recombinant baculovirus is then used to infect susceptible insects or cultured insect cells and the protein product from the incorporated selected gene is produced from the infection.

Many recombinant baculovirus expression vectors and shuttle vectors are on deposit at the ATCC or the Agricultural Research Culture Collection (Peoria, Ill.).

In this case, the corrected full length cDNA insert was subcloned into the pBlueBac transfer vector (see J. Vialard *et al.*, J. Virology 64:37 1990) under the polyhedron promoter and cotransfected into *Spodoptera frugiperda* cells ("Sf9 cells") with wild type baculovirus DNA. (Sf9 cells were either grown in T150 tissue culture flasks in TNM-PH media or as suspension cultures in EX-CELL 401 (JRH Biosciences) in shaker flasks.) The advantage of selecting the pBlueBac transfer vector (commercially available from Invitrogen Corp., San Diego, CA.) is the coexpression of the β -galactosidase protein from the ETL promoter allowing easy selection of recombinant baculoviral plaques in the presence of X-gal.

To construct the pBlueBac transfer vector, part of the E1-19 cDNA clone was first subcloned into M13mp19 followed by site-directed mutagenesis (as noted above) to correct a frame-shift mutation. This insert was re-sequenced and ligated into E1-19 in the pUC18 vector. The full length cDNA was constructed in a pT7-7 vector (see S. Tabor and C.C. Richardson, Proc. Natl. Acad. Sci. USA 82:1074 1985) by ligation of the corrected E1-19 clone with E1-14a at the

-21-

unique SalI site. The 5' NcoI site of E1-19 at the initiation ATG codon was filled in using Klenow fragment and ligated to EcoRI linkers for insertion into the pT7-7 vector. This clone, designated pT7/HDP α , was restricted with EcoRI and filled in with Klenow. This DNA was then digested with DraI to remove most of the 3' untranslated region, and ligated with XbaI linker followed by subsequent restriction with XbaI and ligated into the unique NheI site in the pBlueBac transfer vector. The 5' end manipulations and site-directed changes of the frame-shift error in E1-19 were confirmed by dideoxy sequencing. The resulting construct was named pBlueBac/HDP α (see Figure 4). This pBlueBac/HDP α transfer vector was then co-transfected with wild type baculovirus DNA into Sf9 cells by CaPO $_4$ transfection as described by M. Summers and G.E. Smith, Bulletin No. 1555, Texas Agriculture Experimentation Station (College Station, TX) (1988). Recombinant AchDP α baculovirus was detected by X-gal in agarose overlays. A second recombinant virus was made by ligating the EcoRI-DraI fragment containing the full length cDNA from pT7/HDP α plasmid into EcoRI-SmaI digested pVL1392 transfer vector. This plasmid, pVL1392/HDP α was also co-transfected with wild type baculovirus DNA into Sf9 cells. The resulting recombinant virus, 1392 α , was detected by staining with neutral red in agarose overlays. A control recombinant baculovirus, Ac β Gal, expressing the E. coli β -galactosidase protein was also made using the pAc360 β Gal transfer plasmid.

Ten occlusion minus blue viral plaques were plaque purified and screened for the presence of human DNA polymerase α by immunoblot analysis using two polyclonal antisera, DPN and DPC, specific for

-22-

peptide sequences at the N- and the C-terminals of the human DNA polymerase α catalytic polypeptide. See K. Hsi *et al.*, Nucleic Acids Res. 18:6231 (1990). After four rounds of plaque purification, 7 of the plaque purified viruses expressed the human polymerase α . The resulting recombinant virus, named AchDP α , expresses the full length recombinant human DNA polymerase α catalytic subunit from its natural ATG start codon under control of the polyhedron promoter. In addition, the full length human DNA polymerase α cDNA was inserted into the pVL1392 transfer vector and recombinant virus isolated after co-transfection into Sf9 cells. This virus, 1392 α , also expresses equivalent amounts of functional recombinant human polymerase α as does the AchDP α virus.

The level of expression and solubility of recombinant DNA polymerase α in Sf9 insect cells has been analyzed. The AchDP α baculovirus infected Sf9 cells were harvested every 12 hours after infection and analyzed for the expression of human polymerase α protein by SDS-PAGE (Figure 5A, 5B, 5C, 5D, and 5E) and for enzymatic activity (Figure 5F) by polymerase assays. The SDS gels were stained by Coomassie Blue (Figures 5A, 5C to 5E). The presence of human DNA polymerase α was verified by immunoblot analysis (Figure 5B) with serum antibodies directed against the N- or the C-terminal peptides of human DNA polymerase α catalytic polypeptide named DPN and DPC, respectively.

In Figures 5A-E, lanes 0, 12, 24, 36, 48, 60, 72, and 84, represent the hours of cell harvest post-infection. Figure 5 is a Coomassie stained gels of whole cell lysates from 3×10^5 Sf9 cells. The arrows indicate the expressed intact 180 kDa recombinant

-23-

polymerase α protein which appears to be the most abundant protein expressed. A second most abundant polypeptide expressed is of 140 kDa. Several proteins in minor quantity ranged from 160 to 105 kDa appeared after 36 to 84 hours post-infection are proteolytically degraded forms of the expressed recombinant polymerase.

Figure 5B is an immunoblot of the whole cells lysates shown in Figure 5A equivalent to 3×10^4 cells, along with polymerase α isolated from human "KB" cells used as standard for comparison by the antisera, DPN and DPC. DPN detects the intact 180 kDa recombinant human polymerase α and also detects several proteolytic species of the recombinant human polymerase α protein of 140, 90, and 50 kDa. DPC detects the intact 180 kDa recombinant protein and proteins of 160 and 140 kDa, and several minor proteolytic species ranged from 105 to 60 kDa.

Figure 5C is a gel showing the electrophoresis of 100 μ g protein from the soluble cell lysates which are equivalent to the amount of protein from 5×10^5 Sf9 cells. The accumulation of a protein of ~105 kDa after 60 hr post-infection detected by neither DPN nor DPC are possible proteolytic degraded recombinant human polymerase α from both the N- and C-termini.

Figure 5D is a gel showing the electrophoresis of polymerase α resolubilized from a high salt extraction. Fifty (50) μ g of protein was loaded onto the gel which is equivalent to the amount of protein resolubilized from 1.6×10^6 cells. Figure 5E is a Coomassie stained gel of the insoluble polymerase pellet from 1.5×10^6 cells.

Figure 5F shows the total DNA polymerase units recovered in the soluble and high salt resolubilized lysates from each time point. (-□-), activity from

-24-

the soluble lysates; (-▲-), activity from the resolubilized lysates by high salt extraction; (-■-), the sum total activities in the soluble and the high salt resolubilized lysates.

5 In general, the data shows that expression of recombinant human polymerase α can be detected in whole cell extracts as early as 24 hours after infection with AcHDP α and reaches a maximum level of expression between 48 and 60 hours post-infection
10 (Figure 5A, 5B and 5F). Therefore, the optimal time for isolation of enzymatically active recombinant polymerase α therefore is 48 hours when Sf9 cells are infected with AcHDP α at a multiplicity of infection (M.O.I.) of 10.

15 The amount of recombinant polymerase α protein expressed in insect cells was quantitated by densitometric analysis of Coomassie blue stained gels. It is expressed at a level of approximately 12% of the total cellular protein. The amount of
20 soluble and enzymatically active recombinant human polymerase α obtained is dependent on the time of harvest as well as on the method used for cell lysis. Cells were lysed either by sonication in isotonic buffer or by hypotonic Dounce homogenization or
25 treatment with nonionic detergents such as Triton X-100 or Nonidet P-40. These methods yield near 50% soluble recombinant polymerase α protein at 48 hours post-infection. To minimize proteolytic degradation and time of manipulation, sonication has been used in
30 most of the experiments described here. About 50% of the expressed polymerase α protein can be isolated in soluble enzymatically active form by sonication of insect cells at 48 hours post-infection in isotonic buffer (Figure 5C). An additional 15-20% of the
35 polymerase α activity could be resolubilized from the

-25-

insoluble pellet by a high salt extraction (Figure 5D). After high salt extraction, the most abundant protein remaining in the insoluble pellet was the recombinant human polymerase α protein which comprised approximately 30-35% of the total expressed recombinant polymerase α protein. After 48 hours of infection, the solubility of the produced recombinant polymerase α protein decreased as post-infection time progressed. At later time points such as 72 and 84 hours post-infection, a much lower amount of undegraded recombinant p180 polymerase α protein is detected in soluble cellular lysates by Coomassie staining (Figure 5C).

Using previously produced polyclonal antisera directed against 20 amino acid residues at the N- and C-termini of human polymerase α catalytic polypeptide which are designated DPN and DPC, respectively, a specific labile site was defined near the N-terminus of the catalytic polypeptide. See K. Hsi *et al.*, Nucleic Acids Res. 18:6231 (1990). To analyze the proteolytic susceptibility of the over-produced recombinant polymerase α , whole cell lysates were transferred to membrane and immunoblotted with antisera, DPN and DPC. Immunoblot analysis indicates substantial degradation of the expressed polymerase α protein from both the N- and C-termini even at 36 hours post-infection (Figure 5B). The predominant protein detected by the antisera was 180 kDa and the most abundant proteolytic product is the p140 detected only by the DPC antibody. The degradations of recombinant protein detected by the DPN antisera have not been observed in polymerase α protein purified from cultured human cells. The degraded polymerase protein may not be detectable in

-26-

preparations from cultured human cell due to the low quantity of polymerase.

Figure 5F illustrates the total enzymatic activity of the soluble recombinant polymerase α in the cell lysates and the activity resolubilized by high salt extraction of the cell pellet. The amount of soluble and assayable recombinant human polymerase α activity reaches a maximal level after 48 hours post-infection. After 48 hours, the soluble activity slightly decreases in the later post-infection harvested cells. As shown in the profile (Figure 5C and 5F), the degraded p180 protein retains nearly full activity. This is in agreement with several previous reports in which demonstrate that degraded forms of the polymerase α protein retain full enzymatic activity. The assayable recombinant DNA polymerase α activity at 48 hours post-infection is approximately 1000 fold over that in uninfected Sf9 cells. This over-production can be improved by infection of Sf21 cells in serum free media. When human polymerase α molecule is quantitated on a per cell basis, the Sf9 cells grown in TNM-FH media produce about 6×10^6 molecules of soluble and catalytically active recombinant enzyme per cell. Sf21 cells produce approximately twice this much recombinant polymerase α per cell. Furthermore, when Sf21 cells are grown in EX-CELL 401 media, they produce four times the amount of recombinant human polymerase α in approximately 2.4×10^7 molecules per cell as compared to Sf9 cells grown in TNH-FH media (data not shown). Comparing the expression of the recombinant human polymerase α in insect cells to those described in transformed or normal cultured human cells, the recombinant human polymerase α is over-produced greater than 1,000-fold. Moreover,

-27-

AchDPA α infected insect cells grown in suspension culture produce comparable levels of recombinant polymerase α activity. The assayable polymerase activity solely represents the recombinant human DNA polymerase α activity and not the endogenous baculovirus DNA polymerase activity, since the baculovirus polymerase gene is transcribed as early as 2 hours post-infection, reaches a maximum level at 6 hours and declines to negligible level by 12 hours post-infection.

III. PROTEIN PURIFICATION AND CHARACTERIZATION

For protein purification, Sf9 cells infected at a multiplicity of infection (M.O.I) of 10 were harvested as early as 36 hours to 55 hours post-infection. Briefly, cells were removed from T150 flasks by shaking and harvested by centrifugation at <200xg. The cells were washed in serum free Grace media and sonicated for 10 seconds in 20% ethylene glycol, 100mM Tris HCl, pH7.5, 100 mM NaCl, 1 mM EDTA, 1mM β -mercaptoethanol, 1 mM phenylmethanesulfonyl fluoride and 1 mM Sodium bisulfite. This extract was centrifuged for 10 minutes at 12,000xg. The supernatant was removed and saved as the soluble extract while the insoluble pellet was extracted with 600 mM NaCl, 50 mM potassium phosphate, pH 7.5, 20% ethylene glycol, 1 mM EDTA, 1 mM β -mercaptoethanol, 1 mM phenylmethanesulfonyl fluoride and 1 mM Sodium bisulfite. This extraction was again centrifuged at 12,000xg for 10 minutes and the supernatant designated as the high salt solubilized fraction. The soluble and high salt solubilized fractions were then combined and adjusted to 100mM ionic strength with 20% ethylene glycol, 1mM EDTA and 1mM β -mercaptoethanol, and batch adsorbed

-28-

onto phospho-cellulose equilibrated in 20% ethylene glycol, 100mM potassium phosphate, pH 7.5, 1mM EDTA and 1mM β -mercaptoethanol. The resin was washed extensively with the equilibration buffer and the enzyme removed by step elution with 300 mM potassium phosphate as described by T. Wang *et al.*, J. Biol. Chem. 259:1854 (1984).

Immunoprecipitation. Immunoprecipitation of the recombinant human DNA polymerase α polypeptide was performed generally as described by S.W. Wong *et al.*, J. Biol. Chem. 261:7958 (1986) and P.A. Fisher and D. Korn, J. Biol. Chem 252:6528 (1977). Briefly, monoclonal SJK-237-71 (ATCC Catalogue # CRL 1645, 6th Ed. 1988), was used to immunoprecipitate antigen proteins from Sf9 cell lysates which were infected with either AchDP α or with a control recombinant baculovirus, Ac β Gal. After separation on SDS-PAGE the gel was stained with Coomassie blue (Figure 6A). Lanes 1 and 2, 100 μ g of 32 P α -labeled AchDP α and Ac β Gal-infected cell soluble lysates, respectively. Lanes 3 and 4 are immunoprecipitations from the AchDP α infected cell soluble lysate, and Ac β Gal-infected cell soluble lysate, respectively. The 55 and 25 kDa peptides in all the immunoprecipitations represent the heavy and light chains.

A densely staining polypeptide of 180 kDa is immunoprecipitated from Sf9 cells infected by AchDP α , (Figure 6A, lane 3), whereas no protein of this size range is immunoprecipitated from the mock infected cell lysates (Figure 6A, lane 4). Because of the large excess of β -galactosidase present in the lysate from the Ac β Gal infected cells, a small amount of β -galactosidase was carried over in the immunoprecipitation. The two neutralizing monoclonal

-29-

antibodies, SJK-132-20 and SJK-287-38 (ATTC Catalogue #CRL 1640 and 1644), are also able to immunoprecipitate the single subunit recombinant polymerase α (data not shown).

5 It has been reported that human polymerase α catalytic polypeptide (p180) and the p70 subunit are phosphoproteins and the phosphoamino acids are phosphoserine and phosphothreonine. Furthermore, it was found that the catalytic subunit p180 is
10 phosphorylated through the cell cycle but hyperphosphorylated during mitotic phase. The p70 subunit is only phosphorylated in mitotic cells. To test whether the single subunit recombinant human polymerase α is phosphorylated in AchDP α infected
15 insect cells, AchDP α infected Sf9 cells were incubated with ^{32}P -orthophosphate 24 hours post-infection in normal TNM-FH media which contains 1.0 g/liter of cold sodium phosphate.

20 Specifically, in vivo phospholabeling of polymerase α in Baculovirus infected cells was accomplished by adding (at 26 hours post-infection) 330 μCi of inorganic ortho- $^{32}\text{PO}_4$ to 1.5×10^7 Sf9 cells in a T150 flask in normal TNM-FH media (1.8 mCi/mmol). Sf9 cells were harvested at 38 hours
25 post-infection, lysed by sonication and immunoprecipitated with SJK237-71 monoclonal antibody covalently linked to Sepharose 4B by mixing the lysate end over end with the Sepharose beads at 4°C for 1 hour. Beads were then washed 10 times with
30 radioimmune precipitation buffer followed by boiling in SDS gel loading dye and loaded directly onto an 8% SDS polyacrylamide gel. After electrophoresis (SDS-PAGE) the gel was then stained with Coomassie brilliant blue, destained, dried and subjected to
35 autoradiography. Figure 6B is the autoradiogram of

-30-

immunoprecipitation of lanes 3 and 4 of Figure 6A (24 hour exposure).

Immunoprecipitation of labeled cell lysates with SJK237-71 demonstrates a readily detectable phosphoprotein of 180 kDa from lysates of AchDP α infected Sf9 cells labeled in the high phosphate medium (Figure 6B, lane 1).

Immunoaffinity Purification. Immunoaffinity purification of the recombinant human DNA polymerase α polypeptide was performed generally as described by T. Wang *et al.*, J. Biol. Chem. 259:1854 (1984). Fifteen-T150 flasks each containing 1.5×10^7 Sf9 cells infected with AchDP α at multiplicity of 10 were harvested at 40 hours post-infection as described above and a crude cytoplasmic extract was prepared. The crude lysate containing the recombinant polymerase α was batch absorbed on phosphocellulose in buffer (1 mM mercaptoethanol, 1mM EDTA and 20% ethylene glycol). After extensive washing, the enzyme was removed by step elution in 0.3 M KPO $_4$ (pH 7.5). Monoclonal antibody SJK237-71 was preadsorbed on a Protein A-sepharose CL-4B column (Sigma Chemical Co.) to make an IgG-Protein A matrix. The phosphocellulose eluate was adjusted to pH 8.2 and added to the matrix. After washing the column with buffer, the retained polymerase activity was eluted with 50 mM Na acetate (pH 5.5) containing 1M KCL. After dialysis, the enzyme was further purified on a denatured calf thymus DNA-cellulose column.

The results of the immunopurification protocol are presented in Table I and the recombinant polymerase α protein profiles for each step of the purification are shown in Figure 7. Figure 7A is a Coomassie stained gel of active polymerase α

-31-

fractions throughout the purification. Lane 1, 120 μ g of soluble crude cell lysate from uninfected Sf9 cells. Lane 2, 120 μ g of soluble crude cell lysate from AchDP α infected Sf9 cells. Lane 3, 20 μ g of the phosphocellulose peak fraction. Lane 4, 50 units of polymerase α activity from the pooled SJK-237-IgG Protein A column eluate. Lane 5, 50 units of polymerase α activity from the pooled DNA-cellulose fractions. Lane 6, 40 units of active four subunit-DNA polymerase α /primase complex from cultured KB cells used as a standard for comparison. Less amount of proteins appears to be loaded in lane 6 than in lane 5 as indicated by the tightly associated monoclonal antibody heavy chain in 1:1 ratio in lane 5 and 6. In lane 6, the catalytic polypeptide of KB polymerase α contains not only the intact 180 kDa species but also the proteolytically degraded forms ranged from 160 to 120 kDa. Comparable specific activity of the recombinant single subunit polymerase α and the four subunit polymerase is estimated according to the combined amount of catalytic polymerase α protein of the KB cells.

Figure 7B is a silver stained gel of immunopurified polymerase α . Lane 1, 5 units of 4 subunit-DNA polymerase α /primase from KB cells from an antigen preparation using covalently linked SJK237-Sepharose 4B. Lane 2, 5 units of active single subunit recombinant polymerase α immunopurified from AchDP α infected Sf9 cells by SJK237-Protein A Sepharose 4B, a similar immunopurified preparation to that shown in lane 5 of Figure 7A. Arrows designate the p180 and p165 polypeptides of the polymerase as well as the heavy chain from the hybridoma SJK237-IgG.

-32-

Table 1: Immunopurification of the recombinant human DNA polymerase α from insect Sf9 cells

5	Active Fraction	Total Protein ^a	Total Units ^b	Sp. Activity ^c	% Yield
	Crude cell lysate	98.4 mg	38,400	390	100
	Phosphocellulose	18.7 mg	35,710	1910	93
10	IgG-Protein A	0.18 mg	13,050	70,000 ^d	34
	DNA Cellulose	38 μ g	7,680	200,000 ^d	20

^a Protein concentrations of crude cell lysates and phosphocellulose fractions were determined by Bradford analysis using BSA as the standard. IgG-protein A and DNA cellulose protein fractions were determined by densitometric analysis of Coomassie blue stained gels containing known amounts of BSA standard.

^b One unit of DNA polymerase is defined as the amount of protein to incorporate one nmole of labeled dNMP per hour.

^c Specific Activity is defined as units DNA polymerase per total mg of protein.

^d Specific activity of non-IgG protein. Specific activity with IgG is 9700 and 57,000 units/mg for IgG-protein A and DNA cellulose fractions, respectively.

-33-

From the data, it is apparent that, in the soluble crude cell lysate, the specific activity of the polymerase α activity is at least 100-fold higher than that obtained by a traditional crude human KB cell extract (compare, for example, with the data in Table I of T. Wang *et al.*, J. Biol. Chem. 259:1854 on page 1856). Subsequent purification increases this specific activity. In particular, immunoaffinity chromatography separates the polymerase from other cellular protein and renders it substantially purified.

Importantly, no protein species of a size corresponding to the baculovirus polymerase (114 kDa) was detected by electrophoresis following the last two purification steps, i.e., the SJK237-71-Protein A fraction and DNA cellulose fraction, (Figure 7A, lanes 4 and 5). This further confirms the species specificity of monoclonal antibody SJK237-71, eliminating the possible cross-reactivity of Sf9 insect cell polymerase in the immunoaffinity purified enzyme fractions. (This species specificity was also demonstrated above in Figure 6A, lane 4 with the immunoprecipitation of mock-infected Sf9 cell control.)

Characterization of Affinity. The four subunit-DNA polymerase α /primase complex was purified from human KB cells as described T. Wang *et al.*, J. Biol. Chem. 259:1854 (1984) and used for comparison to the single polypeptide recombinant polymerase α for their respective affinities for dNTP, primer-terminus, and gapped DNA. Prior to the comparison, both the immunopurified recombinant single subunit-polymerase α and the four subunit-polymerase α /primase complex were stored in buffer containing 30% sucrose, 20%

-34-

ethylene glycol, 50mM Tris HCL, pH 8.6, 1mM β ME, 1mM EDTA at -80°C or stored on packed ice at 4°C .

5 The standard assay for DNA polymerase α with gapped DNA was performed according to P.A. Fisher and D. Korn, J. Biol. Chem 252:6528 1977). Reactions were performed using optimally gapped salmon sperm DNA in 20mM Tris·HCl, pH 8.0, 2 mM β -mercaptoethanol ("βME"), 200 $\mu\text{g/ml}$ BSA, 10mM MgCl_2 , 50 μM dNTP's with [α - ^{32}P]dATP as the label. One unit of DNA polymerase 10 is defined as the amount of polymerase that incorporates 1 nmole of labeled dNTP into acid-insoluble DNA at 37°C in 60 min. K_m values for primer terminus were performed on oligo(dT)₁₂: poly(dA)₂₉₀ where an average of five oligodT molecules were 15 annealed per polydA molecule and reaction was performed in 20 mM Tris·HCl, pH 8.0, 2mM β ME, 200 $\mu\text{g/ml}$ BSA, 2mM MgCl_2 and 50 μM [α - ^{32}P]dTTP. All Kinetic parameters were calculated from Lineweaver-Burk plots by the method of least squares. K_m was 20 calculated of the basis of 3'-OH primer termini.

The kinetic parameters determined for dNTPs, primer-terminus and gapped DNA are summarized in Table II and demonstrate no apparent differences 25 between the two forms of DNA polymerase α . The rate of catalysis of these two forms was also measured. The K_{cat} values determined for the two forms of polymerase α are of a similar order of magnitude.

Reactivity to Aphidicolin and N^2 -(p-n-butylphenyl)-dGTP. DNA polymerase α is distinct from 30 DNA polymerases δ or ϵ for its sensitivity to the dNTP analog, N^2 -(p-n-butylphenyl)-dGTP. This compound is a potent inhibitor of the DNA synthetic capacity of DNA polymerase α , but not of DNA polymerases δ or ϵ . Another compound, aphidicolin, a potent DNA

-35-

synthesis inhibitor in vivo, inhibits all three DNA polymerases in vitro. The recombinant single subunit-polymerase α and the four subunit-polymerase α /primase complex from cultured human KB cells were comparatively assayed in the presence of increasing concentrations of N^2 -(p-n-butylphenyl)-dGTP or aphidicolin. These inhibitor studies were performed as standard DNA polymerase assays with the amount of added inhibitors as following: N^2 -(p-n-butylphenyl)-dGTP inhibition reactions were performed in the concentration range of 0.1 to 50 μ M, while the aphidicolin inhibition reactions were performed in the concentration range of 1 to 1000 μ M. Inhibition curves were plotted and data presented as concentration of each inhibitor which causes 50% inhibition. Both forms of polymerase α were extremely sensitive to these compounds and their levels of sensitivity were identical, Table II.

Thermosensitivity. The thermostability of the two forms of polymerase α was also compared; this was done by preincubation at 37°C for various times before assaying polymerase activity. Both forms of polymerase α activity were found to decay at a nearly identical rate to ~67% of the original activity after 30 minutes at 37°C, Table II. Moreover, activities of both forms of the polymerase α also were found to decrease at approximately the same rate when stored at 4°C (data not shown). These decreases in polymerase activity was not due to proteolysis of the polymerase protein. Gel analysis of the two forms of enzyme after prolonged storage at 4°C or after incubation at 37°C demonstrates only nominal degradation of the 180 kDa catalytic polypeptide (data not shown).

-36-

Table II: Properties of the four subunit and single subunit recombinant DNA Polymerase α

5	Property*	Enzyme	
		4 subunit Pol α	Single subunit recombinant Pol α
	K_m (dNTP), μM	1.55	1.20
10	K_m (primer terminus), μM	0.3	0.4
	K_m (DNA in nucleotide), mM	0.22	0.19
	K_{cat} sec ⁻¹	1.3	1.6
15	50% BuPdGTP inhibition, μM	0.22	0.22
	50% Aphidicolin inhibition, μM	13	20
20	Thermostability, % activity after 30 min at 37° C	66	68

25 Reactions were performed using optimally gapped salmon sperm DNA in 20mM Tris·HCl, pH 8.0, 2 mM β ME, 200 $\mu g/ml$ BSA, 10mM MgCl₂, 50 μM dNTP's with [α -³²P]dATP as the label. K_m values for primer terminus were performed on oligo(dT)₁₂: poly(dA)₂₉₀ in 20 mM Tris·HCl, pH 8.0, 2mM β ME, 200 $\mu g/ml$ BSA, 2mM MgCl₂ and 50 μM [α -³²P]dTTP.

30 All Kinetic parameters were calculated from Lineweaver-Burk plots by the method of least squares.

-37-

Absence of Associated Exonuclease Activity.

Purified four subunit-DNA polymerase α /primase complex from a variety of species does not contain detectable 3'-5' exonuclease proofreading activity. It has been reported that a cryptic proofreading 3'-5' exonuclease is present in the *Drosophila* DNA polymerase α catalytic subunit when separated from the other associated subunits. The over-produced single subunit recombinant human polymerase α from ACHDPA infected Sf9 cells provides an ideal enzyme to investigate the presence of a cryptic exonuclease in the polymerase α from somatic human cells.

Proofreading exonuclease activity of the recombinant single subunit polymerase α was assayed with a singly primed M13mp18 template, primed either with a matched 24mer (Figure 8B) or a mismatched 29mer (Figure 8A), each ^{32}P -labeled at the 5' end by T4 polynucleotide kinase. The mismatched 29mer contains 9 mismatched T's on the 3' terminus. Correct proofreading of this 29mer annealed to M13mp18 would produce a 20mer in the absence of deoxynucleotide triphosphates (dNTPs).

To perform the assay, two oligonucleotides, the universal primer, a 24mer (matched primer) and RD29mer (mismatched primer), were 5'-end labeled with ^{32}P -ATP and annealed to M13mp18 single stranded template. 0.1 pmole of this primed M13 was incubated in 20 mM Tris-HCl, pH 7.5, 10 mM MgCl_2 , 20 mM KCl and 2 mM dithiothreitol at 37°C in the presence of polymerase in a final volume of 12 μl . Aliquots of 3 μl each were removed at 0, 2.5, 10 and 30 minutes into an equal volume of deionized formamide containing 1mM EDTA, 0.1% xylene cyanol and 0.1% bromophenol blue and placed on ice. After heating at 95°C for 5 minutes, one-half of the sample was loaded

-38-

onto an 18% polyacrylamide, 7 M urea, Tris-borate-EDTA gel.

5 The recombinant single subunit polymerase α of the present invention was assayed and compared to the four subunit-polymerase α /primase complex, human DNA polymerase ϵ , and T4 DNA polymerase, Figure 8. In both Figures (A) and (B) lanes 1 represents the incubation of the primed M13 with no enzyme for 2.5 minutes; lanes 2, 3, and 4 are the incubations with the primed M13 substrate for 2.5, 10, and 30 minutes at 37°C in the presence of 1.0 units of the four subunit-polymerase α /primase complex from KB cells, respectively; lanes 5, 6, and 7, are incubations with 1.0 units of the single subunit recombinant polymerase α immunopurified from AchDP α infected Sf9 cells for 2.5, 10, and 30 minutes, respectively; lanes 8, 9, and 10, are incubations in the presence of 0.4 units of purified HeLa cell DNA polymerase ϵ for 2.5, 10, and 30 minutes, respectively; and lanes 11, 12, and 13 are incubations in the presence of 0.02 units of phage T4 DNA polymerase for 2.5, 10, and 30 minutes, respectively.

25 It is clear that after 30 minutes of incubation, no apparent 3'-5' exonuclease was detected in either the recombinant single subunit-polymerase α or the four subunit-polymerase α /primase complex assayed with either the mismatched primer-template (the 29mer; Figure 8A) or the matched primer-template (the 24mer; Figure 8B). (The matched primer contains a small amount of contaminating primer of 19 bases in length and is not the result of exonuclease activity as seen in the control in lane 1.) In contrast, both the polymerase ϵ from HeLa cells and T4 DNA polymerase digested both the mismatched and the matched primer-templates but with a specificity for

35

-39-

the mispaired primer. These results demonstrate that, unlike the finding of a cryptic exonuclease in *Drosophila* embryo polymerase α , the catalytic polypeptide of human DNA polymerase α either in the four-subunit complex form or as a single-subunit lacks detectable proofreading exonuclease.

A recent report describes the purification and reconstitution of the yeast polymerase α catalytic subunit with the p86 subunit. The exonuclease activity of the four subunit yeast polymerase α /primase complex versus the single catalytic subunit yeast polymerase α and the reconstituted p180-p86 complex were investigated. See R.G. Brooke *et al.* J. Biol. Chem 266:3005 (1991). No proofreading activity was detected in any of the yeast polymerase forms, but a 3'-5' exonuclease activity was detected using substrates such as poly(T)₆₀₀·[³²P]dCMP_{0.4} and on longer polynucleotides but not with short polynucleotides such as poly(dT)₂₅ or poly(dT)₅₀ as substrate.

To test whether the recombinant human single subunit of the present invention could release label from this kind of nonphysiological synthetic substrate, poly(dT)₂₀₀₀ was end-labeled with [α -³²P]dCTP by calf thymus terminal deoxynucleotide transferase. Possible 3'-5' exonuclease activity was tested using the poly(dT)₂₀₀₀[³²P]dCMP_{0.5} as substrate in the nucleotide release assay described by Brooke *et al.* After 30 minutes at 30°C 35 μ l of the 40 μ l reaction was removed and acid precipitated with carrier DNA as performed in the standard DNA polymerase assay. The remaining 5 μ l of reaction was mixed with an equal volume of 95% formamide sequencing loading dye of which 2 μ l was loaded onto a 7 M urea, 8% polyacrylamide sequencing gel followed by electrophoresis and autoradiography. Figure 9 shows

-40-

the results from this assay. Incubation of substrate with 4 subunit-polymerase α /primase complex, (lane 1); recombinant single subunit-polymerase α , (lane 2); recombinant human polymerase β , (lane 3); T4 DNA polymerase, (lane 4); and buffer only controls, (lanes 5 and 6). The released dCMP label is designated by the lower arrow and the poly(dT)₂₀₀₀·[³²P]dCMP₀, substrate is designated by the upper arrow.

The results of Figure 9 indicate that both the four subunit and single subunit recombinant polymerase α do not contain any detectable exonuclease activity on this nonphysiological synthetic substrate. As control for the assay, the recombinant human polymerase β and T4 DNA polymerase were used as negative and positive controls, respectively. As with the native human polymerase β , no detectable exonuclease activity was found with the recombinant human polymerase α of the present invention, while the T4 DNA polymerase released all of the dCMP label.

Aliquots of the reaction were also quantitated by acid precipitation and counted. The exclusion of exonuclease activity from polymerase activity in recombinant human single subunit and four subunit polymerase α /primase complex as well as the recombinant human DNA polymerase β demonstrate $<3 \times 10^{-6}$ nuclease/polymerase activity. In contrast, the T4 DNA polymerase released $>99\%$ of the label. This quantitative analysis makes it clear that the human recombinant single subunit polymerase α is devoid of any detectable 3'-5' exonuclease activity.

-41-

IV. TEMPLATE-DEPENDENT ENZYMATIC SYNTHESIS

Enzymatic synthesis that involves nucleic acid, either solely as a template (e.g., translation involves the use of nucleic acid as a template to make polypeptides) or as both a template and a product (replication and transcription use nucleic acid as a template to produce nucleic acid) is hereinafter referred to as "template-dependent enzymatic synthesis."

In the case of replication, nucleic acid polymerases replicate a nucleic acid molecule ("template") to yield a complementary ("daughter") nucleic acid molecule. For example, DNA polymerase I, isolated from E. Coli, catalyzes the addition of deoxyribonucleoside triphosphates to the 3' end of a short segment of DNA ("primer") hybridized to a template strand to yield a daughter of the template, starting from a mixture of precursor nucleotides (dATP, dGTP, dCTP, and dTTP).

This 5' to 3' template-dependent enzymatic synthesis is also called "primer extension." Importantly, the reaction will not take place in the absence of template and primer.

While all DNA polymerases require a 3'-hydroxyl terminus of a preexisting primer for reaction, DNA polymerase α (in its native form) is the only eukaryotic polymerase with a tightly associated primase. See T. Wang, Ann. Rev. Biochem 60:5413 (1991). A "primase" is a class of enzymes capable of accomplishing physiologically significant de novo primer synthesis. T. Wang et al., J. Biol. Chem. 259:1854 (1984).

When the single subunit recombinant human polymerase α of the present invention was compared with the four subunit DNA polymerase α /primase

-42-

complex obtained by traditional methods (see Table II above), the kinetic parameters of the two forms of the polymerase appeared indistinguishable, suggesting that the absence of the other subunits was not critical to these functions.

The issue of the absence of the other subunits has been further examined. The single subunit recombinant polymerase α of the present invention was tested for its processivity and DNA synthetic fidelity.

DNA Synthetic Processivity. Enzymes that synthesize polymers may dissociate after each catalytic event, i.e., they may be "nonprocessive." On the other hand, they may remain bound to the polymer until many cycles of reaction are completed, i.e., they may be "processive." See A. Kornberg, DNA Replication (Freeman and Co. 1980). It is known that the four subunit DNA polymerase α /primase complex is moderately processive, polymerizing 10-20 bases per binding event. See T. Wang, Ann Rev. Biochem. 60:513 (1991).

The processivity of the polymerase subunit of the present invention was determined on singly primed M13mp18 ssDNA in the absence or presence of E. coli single stranded DNA binding protein (SSB) or on oligo(dT) primed poly(dA). For each reaction, 0.6 pmole of [32 P]-5' end labeled singly primed M13mp18 DNA where only one primer molecule was annealed per M13 template molecule was incubated in 50 μ l with 0.015 pmole of designated DNA polymerase in 20 mM Tris-HCl, pH 8.0, 1mM dithiothreitol, 200 μ g/ml BSA, 10 mM MgCl₂, 100 μ M each of dGTP, dCTP, dATP, and dTTP, at 37°C. E. coli SSB was present at a ratio of 8:1 nucleotides:SSB monomer.

-43-

Processivity on Poly(dA) was performed according to J. Syvaoja and S. Linn J. Biol. Chem 264:2489 (1989). Forty six pmole (in molecules) of oligo(dT)₁₂:poly(dA)₂₉₀, where only one oligo(dT)₁₂ molecule was annealed on average per one poly(dA)₂₉₀ molecule, was incubated with 1.1 fmol of DNA polymerase in a 40 μ l volume with 50 μ M [³²P]dTTP (40,000 cpm/pmol) in 2 mM MgCl₂, 20mM Tris·HCl, pH 8.0, 200 μ g/ml BSA and 1mM dithiothreitol at 37°C. For both the M13 and poly(dA) processivity reactions aliquots were removed at the indicated times indicated in the figure legends, phenol:chloroform extracted and ethanol precipitated with 10 μ g salmon sperm DNA as carrier. Reaction samples were resuspended in deionized formamide containing 1mM EDTA, 0.1% xylene cyanole and 0.1% bromophenol blue, and one-half of the sample loaded onto a 8% polyacrylamide, 7 M urea, Tris-borate-EDTA gel.

The DNA synthetic processivity of the single subunit recombinant human polymerase α of the present invention, using two kinds of primer-templates, is shown in Figure 10. Processivity was measured by using singly primed M13 template in 40 to 1 excess molar ratio of primer-template to polymerase (Figure 10A). ³²P-5'-end labeled 17mer (0.6 pmole) annealed to 0.6 pmole M13mp18 single stranded DNA was extended by 0.015 pmole of polymerase for 2.5, 10, and 30 minutes. Samples were subjected to electrophoresis and autoradiography. Lane 0, incubation of the singly primed M13 in the absence of any added polymerase. Lanes 1, 2, and 3, represent incubation with the indicated polymerase for 2.5, 10 and 30 minutes, respectively. KB designates the four subunit human KB DNA polymerase α /primase complex, AchDP α designates the single subunit recombinant

-44-

polymerase α from AchDP α infected Sf9 cells, and T4 designates the T4 DNA polymerase. Molecular weight markers are labeled HaeIII-digested ϕ X174 DNA, 29mer and 17mer. The arrow indicates a pause site by the polymerases, corresponding to the lacZ operator sequence located 101 bases from the universal 17mer annealing site.

The single subunit recombinant polymerase α and the four subunit-polymerase α /primase complex exhibited similar DNA synthetic processivity on singly primed M13 (Figure 10A). The presence of secondary structure in the M13 has reproducibly found to cause pausing for the polymerases. Synthesis by all the polymerases tested appeared to pause at the lacZ operator structure as indicated by the arrow in Figure 10A. E. coli SSB could relieve this pause for the T4 DNA polymerase but had a slight inhibitory effect for the single and four subunit DNA polymerase α (data not shown).

The singly primed M13 substrate, the bulk of which is single stranded DNA, has been documented to cause non-productive binding and inhibition to the polymerase α synthetic ability. Thus, DNA synthetic processivity of the two forms of DNA polymerase α were further evaluated on oligodT primed polydA as primer-template in >40,000 to 1 excess molar ratio of primer-template to polymerase (Figure 10B). The recombinant single subunit polymerase and four subunit polymerase α /primase complex and T4 DNA polymerase, 1.1 fmol of each, were separately incubated with 46 pmol oligo(dT)₁₂·poly(dA)₂₉₀ at a primer-template ratio of 1:1, respectively, in 40 μ l volume with 50 μ M [³²P]-dTTP (40,000 cpm/pmol) in 2mM MgCl₂ for 4, 12.5, and 30 minutes. Lanes 1, 2, and 3, incubations with KB four subunit polymerase α /primase

-45-

complex, lanes 4, 5, and 6, incubations with the single recombinant polymerase α and lanes 7, 8, and 9, incubations with T4 DNA polymerase. The molecular weight marker is derived from a dideoxy sequencing reaction. The primer length of 12 has been subtracted from the molecular weight marker to reflect the nucleotides synthesized.

Like the results obtained when using singly primed M13 DNA, both forms of the DNA polymerase α exhibited near identical DNA synthetic processivity on this primer-template. Moreover, the products synthesized by the single subunit and four subunit polymerase α /primase complex on this primer-template demonstrate only an increase in net quantity of the same length with increasing incubation time which reflects a true measure of the processivity. The average length products synthesized by the single subunit and four subunit polymerase α were between 7 and 13 nucleotides which is in excellent agreement with the previous published data. These results demonstrate that both the single subunit-recombinant polymerase α and the four subunit polymerase α /primase complex synthesize DNA in a similarly moderate processivity.

DNA Synthetic Fidelity. The fidelity of the single catalytic subunit and the four subunit polymerase α /primase was measured and compared. Three methods were used to measure a single round of gap filling synthesis within the lacZ α -complementing gene in M13mp2 DNA. See T.A. Kunkel et al. Mol. Cell. Biol. 9:4447 (1989). In all three assays, a gapped, double-stranded M13mp2 DNA is constructed which contains a single-stranded gap as the mutation target. The three assays are: (A) forward mutation

-46-

assay, (B) opal codon reversion assay, and (C) assay for -1 base frame-shift.

5 The DNA synthetic mutational frequency and the error rate of the two forms of DNA polymerase α measured by these three methods are summarized in Table III. The values for mutational frequency and error rate as measured by all three assays for the single subunit recombinant polymerase α and the four subunit polymerase α /primase complex are comparable if not identical. The DNA synthetic fidelity values obtained are also in agreement with the values previously determined for the four subunit polymerase α . These results further support the conclusion that proofreading 3'-5' exonuclease activity is absent in 10 both the single subunit recombinant polymerase α and the four subunit polymerase α /primase complex (Figure 8).

V. CHEMOTHERAPEUTIC SCREENING

20 The recombinant human polymerase α of the present invention is particularly useful for screening chemotherapeutics for potential mutagenicity and carcinogenicity. In one embodiment, recombinant human DNA polymerase α is employed to test for incorporation of analogs of the normal 25 nucleotide bases during DNA replication.

One analog of a normal nucleotide base that has been of particular interest lately is 3'-azido-thymidine ("AZT"). AZT was prepared in 1978 by Prusoff and T.S. Lin at Yale University. It has been 30 found active as an antiviral, exclusively against retroviruses. Samuel Broder and Robert C. Gallo of the National Cancer Institute and scientists from Wellcome Research have found that AZT blocks in vitro cytopathic effects of the AIDS virus; 3'-azido-

-47-

Table III: Fidelity of AChDPA single catalytic polypeptide and KB 4 subunit polymerase α /primase

5	DNA Polymerase	Plaques scored		Mutant frequency	Error rate ¹
		total	mutant		
	<u>Forward mutation assay</u>	(light blue and colorless)		($\times 10^{-4}$)	
	Single catalytic polypeptide	4,966	148	300	
	4 subunit polymerase/primase	5,119	145	280	
10	<u>Base-substitution reversion assay</u>		(blue)	($\times 10^{-6}$)	
	Single catalytic polypeptide	420,000	154	370	1/4900
	4 subunit polymerase/primase	420,000	157	370	1/4900
	<u>Minus-one frame-shift reversion assay</u> ²		(blue)	($\times 10^{-5}$)	
	Single catalytic polypeptide	750,000	1243	170	1/1800
15	4 subunit polymerase/primase	730,000	1122	150	1/2200

The background mutant frequency for uncopied DNA was 6.7×10^{-4} for the forward mutation assay, 2×10^{-6} for the base-substitution reversion assay and 1×10^{-5} for the frame-shift reversion assay.

20

¹ Error rate is calculated as described in reference (25).

² The percentage of light blue of total blue plaques was 9%. The light blue plaques represent nonreiterated base frame-shifts and dark blue plaques represent -T errors in the TTTT run.

-48-

thymidine 5'-triphosphate apparently inhibits the viral reverse transcriptase, thereby inhibiting viral replication.

5 To test the usefulness of the recombinant human DNA polymerase α of the present invention as a reagent to screening analogs, the recombinant human DNA polymerase α enzyme was purified from Sf9 cells infected with the AchDP α recombinant baculovirus and the purified polymerase α enzyme was used to test the
10 incorporation of AZT into DNA. Briefly, a running start and standing start primer were separately 5' end labeled with [^{32}P] and annealed to their respective templates as shown below:

Standing start primer-template pair:

15 5'- ^{32}P -TGA CCA TGT AAC AGA GAG-3'
3'-ACT GGT ACA TTG TCT CTC ATT CTC TCT CTC TTC TCT-5'

Running start primer-template pair:

5- ^{32}P -CGC CCA GCG GGC AGA G-3'
3'-GCG GGT CGC CCG TCT CTT ACC TCT TCT CTC CTC TTC TCT-5'

20 These primer-templates were used to test for the incorporation of either dTTP or AZTTP at the site complementary to the outlined "A". The standing start primer-template was extended by the 0.07 units recombinant polymerase α using either dTTP or AZTTP
25 over a time course from 0 to 10 minutes. Following the reaction, the samples were subjected to denaturing gel electrophoresis and autoradiography. Incorporation by either dTTP or AZTTP extends the original primer length of 18 bases to 19 bases (see
30 arrows in Figure 11A).

Similarly, the running start primer-template was extended by the 0.09 units recombinant polymerase α

-49-

in the presence of dATP and either dTTP or AZTTP over a time course from 0 to 10 minutes. Incorporation by AZTTP extends the original primer length of 16 bases to 19 bases (see arrows in Figure 11B).

5 The data of Figure 11 indicates that the recombinant human DNA polymerase α of the present invention is useful as a reagent for screening analogs. This finding that the anti-viral agent AZT is incorporated by the human replication machinery is
10 predictive of mutations in humans.

VI. DRUG DESIGN

The previous section illustrates how the present invention is particularly useful for screening chemotherapeutics for potential mutagenicity and
15 carcinogenicity. It is also contemplated, however, that the present invention be used to design drugs, including drugs with polymerase inactivation properties.

For example, while AZT appears to be the drug of
20 choice at this time for treating AIDS, some results indicate that AZT exhibits toxicity in a clinical setting (in addition to its potential mutagenicity shown by the data described above). Clearly, there remains a strong need for new antiviral agents,
25 especially those with low toxicity to normal cells.

Intensive efforts to develop therapies which can prevent or block the development of serious clinical symptoms in AIDS patients are under way. For the most part, these efforts have focused on the use of
30 nucleotide analogue drugs which inhibit reverse transcriptase. See e.g., U.S. Patent No. 4,916,122 by Chu et al., hereby incorporated by reference. The goal is to find drugs which are more selective and demonstrate greater specificity.

-50-

Traditionally, the search for new drugs capable of interacting with a particular biomolecule, such as a retroviral enzyme, has been somewhat random. The polymerase of the present invention, by contrast, allows for drug design using the knowledge of specific characteristics of the biomolecule as a starting point.

One specific characteristic of the human polymerase α catalytic polypeptide that has heretofore not been known is its detailed structure. This is because, prior to this invention, sufficient amounts of pure polymerase have not been available. By virtue of the present invention, sufficient amounts of the catalytic polypeptide are present to perform analytical work, including x-ray crystallography.

The design of compounds that interact preferentially with, for example, a viral polymerase and not with the polymerase of the present invention can be developed using computer analysis of the three-dimensional structures. Using a set of coordinates for each enzyme, a computer program, and a compound database, putative specific-binding compounds can be identified based on a simple function of interatomic distances.

The interatomic distances can themselves be previously determined by a number of methods known in the art. For example, two-dimensional homonuclear correlated spectroscopy (COSY) generally is the first 2D experiment to be used in analyzing a protein. For those skilled in the art with one-dimensional NMR spectroscopy, COSY provides the kind of information available from a single-frequency decoupling experiment, e.g., which spins are scalar coupled to one another. In a COSY plot, the 1D spectrum lies

-51-

along the diagonal, and the off-diagonal elements are present at the intersection of chemical shifts of groups that are J coupled. The "fingerprint" region contains ($^1\text{H}^N$, $^1\text{H}^\alpha$) cross-peaks from the peptide backbone. The degree of resolution of the "fingerprint" region of a COSY map obtained in H_2O is a good predictor of the success of sequence-specific assignments to be obtained without recourse to isotopic labeling.

Transferred nuclear Overhauser effect (TRNOE) spectra (^1H NMR) relies on different 2D NOE spectra, and, in essence, looks at the conformation of the ligand just after it has dissociated from the protein. The use of TRNOE presumes, however, that the bound and free ligands are in fast exchange on the chemical shift time scale which translates to a ligand K_D greater than or equal to about 10^{-4} M. TRNOE methods are useful to cross-check and augment the distance information obtained by other approaches.

It is not intended that the present invention be limited by the particular method used to obtain structural information. Furthermore, it is not intended that the present invention be limited to a search for any one type of drug; one or more of the molecules may be naturally occurring or may be synthetic. If synthetic, they may be, for example, drug-receptor complexes. If naturally occurring they may or may not be biomolecules ("biomolecules" are herein defined as molecules found in a living organism). If biomolecules, they may be, for example, enzyme-substrate or enzyme-inhibitor complexes.

Finally, it is not intended that the drug design always involve a comparison with another polymerase.

-52-

For example, it may be desired that nucleic acid binding drugs be developed. In such a case, a predictive analysis of the potential impact on the polymerase of the present invention may rely on structural information concerning the polymerase-DNA complex. Such data can be obtained in many ways. For example, De Jong et al. studied the interaction of E. coli phage DNA binding protein with single-stranded DNA. E.A.M. De Jong et al., J. Mag. Res. 80:197 (1988). The technique utilized spin-labeled oligonucleotides; the dipolar interaction between the free electron spin of the spin-label on the substrate causes an increase in the relaxation rate of nearby protons. Spin labels such as TEMPO are good to about 15 angstroms.

VII. VIRAL PROTEIN BINDING

The present invention further contemplates the use of recombinant human DNA polymerase α to test for the binding of viral proteins. In one embodiment, the present invention contemplates co-infection of cells with two expression vectors, one vector coding for the viral protein of interest and the other vector coding for human DNA polymerase α .

As an example of recombinant human DNA polymerase α interaction with viral proteins, the SV40 virus large T antigen was specifically examined. The SV40 virus large T antigen has a wide range of functions, one of which is the ability to transform a permissive cell to the cancerous state. It has already been established that the SV40 large T antigen protein can bind the catalytic subunit of the DNA polymerase α in vitro. See Dornreiter et al. EMBO J. 9:3329 (1990).

-53-

To detect this binding in vivo, baculoviruses were constructed that express both the large T antigen and the human DNA polymerase α catalytic subunit of the present invention in insect cells.

5 For this purpose the 941T baculovirus was employed. See R. Lanford Virol. 167:72 (1988). This expresses the SV40 large T antigen.

The baculoviruses 941T and AchDP α were singly and co-infected into insect Sf9 cells. The cells

10 were incubated for 44 hours, harvested, lysed and proteins immunoprecipitated with either monoclonal antibodies against the human polymerase α catalytic subunit (using monoclonal SJK237-71) or against the large T antigen (using Pab101).

15 Figure 12 shows both the strategy and results of such an experiment. Figure 12A is a Coomassie Blue stained gel of the lysates and anti-Pol α immunoprecipitations. "T" designates the lysate and immunoprecipitation from 941T infected insect cells,

20 "a" from AchDP α infected cells, and "T/a" from the co-infection insect cells demonstrates the association of a protein with identical mass as that of the SV40 T antigen. This protein was confirmed to be the T antigen by a (Western) immunoblot assay

25 (Figure 12B). The T antigen associated with Pol α was detected with the anti-T antigen ("anti-Tag") monoclonal, Pab101.

The reverse experiment was also performed (see right hand panel of Figure 12B). The anti-T antigen

30 monoclonal was used to immunoprecipitate and the anti-polymerase monoclonal was as shown of the right side of the Western blot.

These experiments demonstrate that expression of the human DNA polymerase α in the baculovirus system

35 provides an amendable method to study the interaction

-54-

of cellular and viral proteins, such as the SV40 T antigen, with the polymerase α catalytic subunit.

It is not intended that the present invention be limited to a particular viral protein. In particular, this binding approach is in no way limited to SV40 T antigen. This approach may be extended to other viral proteins.

By way of a further example, the present invention contemplates binding to papillomaviruse proteins. The papillomaviruses are small DNA viruses that induce benign proliferative squamous epithelial and fibroepithelial lesions in their natural hosts. The Bovine papillomavirus type 1 (BPV-1) has served as the prototype for the genetic analysis of the papillomavirus functions. Transformation of rodent cells by BPV-1 has enabled functions important for viral transformation, replication, and transcription regulation to be mapped.

Such studies have revealed that products encoded by the BPV-1 E5 and E6 genes are required for full transformation, E1 products are necessary for viral DNA replication, and E2 polypeptides function both in replication and transcription regulation. To understand BPV-1 replication, Botchan and co-workers have developed an in vitro replication assay. See L. Yang, et al., Nature, in press, (1991). This assay involves adding purified BPV-1 E1 and E2 proteins, a DNA plasmid containing the BPV-1 origin of replication, and radiolabeled nucleotides to murine or human cell extracts and assaying for replication of the plasmid.

The present invention contemplates involvement of the DNA polymerase α complex. Demonstration of this involvement in replication is achieved using the neutralizing antibody SJK-132-20. SJK-132-20 was

-55-

able to abolish in vitro BPV-1 replication while the non-neutralizing antibody, SJK237-71, had no affect on this replication. This experiment suggests that either E1 or E2 or the E1/E2 complex is able to bind and sequester the DNA polymerase α complex at the BPV-1 origin of replication. In this manner, the polymerase of the present invention is useful to map the interaction of the BPV-1 E1 and/or E2 proteins.

From the above it is evident that the present invention provides polymerase α that is functional, and yet free of contaminating protein typically associated with human DNA polymerase α purified by traditional biochemical isolation techniques. This reagent is useful for, among other things, the screening of chemotherapeutics for mutagenicity, particularly where mutations are caused by the incorporation of analogs of the normal nucleotide bases during DNA replication.

-56-

CLAIMS

1. A purified and isolated nucleic acid sequence encoding a human polymerase α catalytic polypeptide.

5 2. A purified and isolated DNA sequence encoding a human polymerase α catalytic polypeptide, said DNA sequence comprising the DNA sequence set out in Figure 3 and its complementary sequence.

10 3. A host cell comprising a functional DNA sequence according to Claim 1 or 2 which renders said host cell capable of expressing a human polymerase α catalytic polypeptide.

4. The host cell of Claim 3, wherein said host cell is from an insect.

15 5. A biologically functional viral vector including a sequence according to Claim 1 or 2.

20 6. A purified and isolated DNA sequence comprising a DNA sequence encoding a polypeptide having an amino acid sequence sufficiently duplicative of that of the human polymerase α catalytic polypeptide to allow possession of the property of processive DNA replication.

7. A DNA sequence according to Claim 6 covalently associated with a detectable label.

25 8. A DNA sequence according to Claim 7, wherein said detectable label is a radiolabel.

-57-

9. A screening method for determining the mutagenicity of chemotherapeutic agents, comprising:

- a) providing a substantially purified human DNA polymerase α catalytic polypeptide;
- b) providing a chemotherapeutic agent suspected of having mutagenic activity;
- c) providing a nucleic acid template for replication;

d) combining said polymerase, said chemotherapeutic agent, and said nucleic acid template under reaction conditions such that, in the absence of said chemotherapeutic agent, said template is replicated.

10. The screening method of Claim 9 wherein said chemotherapeutic is an analog of a normal nucleotide base.

11. The screening method of Claim 10, wherein said analog is incorporated into a polynucleotide during DNA replication.

12. The screening method of Claim 11, further comprising the step of measuring the amount of said analog incorporated into said polynucleotide.

-58-

13. A drug design assay, comprising:

a) providing a substantially purified human DNA polymerase α catalytic polypeptide;

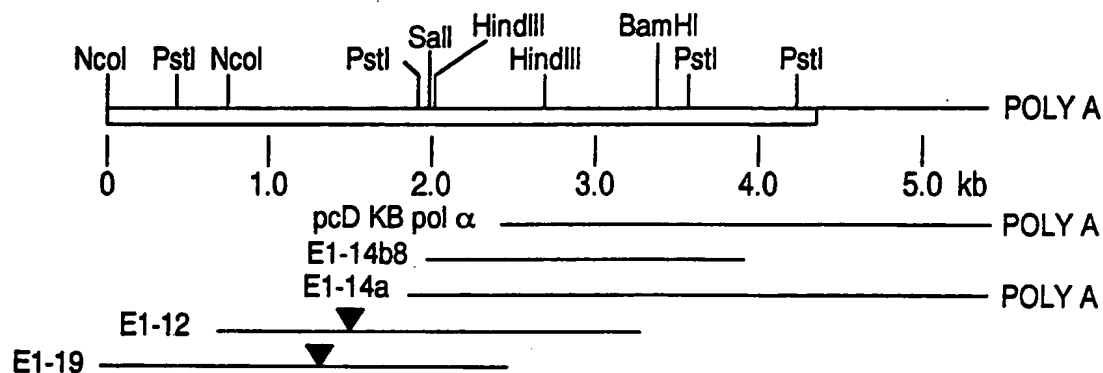
b) providing one or more non-human nucleic acid polymerases;

c) providing a drug with potential polymerase inactivation properties; and

d) comparing the inactivation activity of said drug to said human polymerase with its activity to said non-human polymerase.

14. The drug design assay of Claim 13, wherein said non-human polymerase is a viral polymerase.

1/23

SUMMATION OF MUTATIONS IN POLYMERASE α cDNA CLONES:E1-19 FRAMESHIFT MUTATION:

NUCLEOTIDE

POSITION 1336: E1-12: 5'-G AAC TAT GCT TTT GAG ATA CCT G-3' *

E1-19: 5'-G AAC TAT GCT TTG AGA TAC CTG -3'

BC2#: 5'-G AAC TAT GCA TTC GAG ATA CCT G-3' *

E1-19 FRAMESHIFT MUTATION:

NUCLEOTIDE

POSITION 1519-1538:

EMBO PAPER: 5'-GTA AAA AAG TCC ACA GCT CTT AAT CAG-3'

V K K S T A L N Q

E1-12: 5'-GTA AAA AAG TCC ACA GCT CTT GAA TCA GCC AGT

V K K S T A L E S A S

E1-19: 5'-GTA AAA AGT CCA CAG CTC TTG AAT CAG-3' *

V K S P Q L L N Q

E1-14b8 MISSENCE MUTATION:

NUCLEOTIDE

POSITION 2526: E1-14b8: GGA (Gly)

E1-14a: GCA (Ala) *

FIG. 1

SUBSTITUTE SHEET

2/23

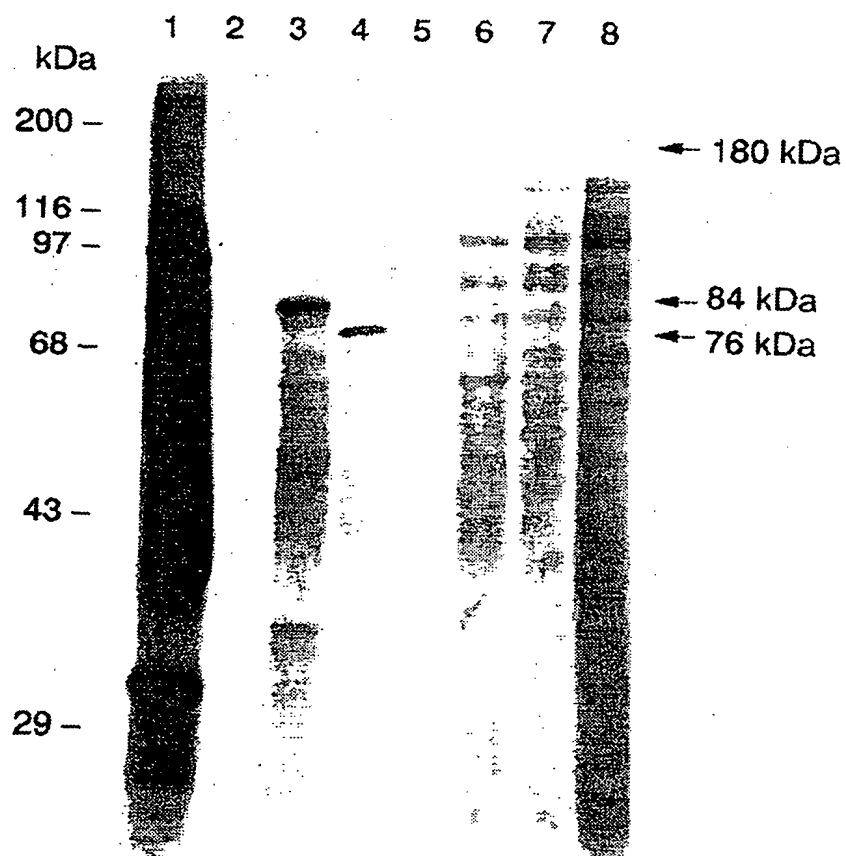


FIG. 2

SUBSTITUTE SHEET

3/23

GGGAGATTGGGACCATGGCACCTGTGCACGGCGAGACTCTCTGTCTGATTCAGGAG 60
TTTGTATCTTCTCGAGCCCGGAGAGAAAAAATCAAGAGGGGCGCCAGAGCCCT 15
FVSSRAARRERKKSKKGRQEL 120
KAAAGACTGAAAAGGCTAAAGCTGGTGAGAGTATAAATATGAAGTCGAGGACTTCAC 35
ERLKKAKAKAGEKKYKYEVEDFT 180
KGTGTTTATGAAGAAGTTGATGAAGACAGTATTCGAAGCTGGTTCAGGCACGCCAGGA 55
GVYEVEDEEQYSSKLVQARQD 240
TATGACTGGATTGTGGATGATGATGGTATGGCTATGTGGAAGATGGCCGAGAGATTT 75
DDWIVDDDGIGYVEDEGREIF 300
TATGATGACCTTGAAGATGATGCCCTTGATGCTGATGAGAAAGAAAGATGGTAAAGC 95
DDLEDDALDADEKKGKDKA 360
KCGAATAAGACAAAGAGGAATGTAAAGAAGCTCGCAGTGACAAACCGAACACATTAA 115
RNKDKRNVNKKLAVTKPNNIK 420
TCAATGTTTCATTGTGTGCTGGAAGAAACTGCAGATAAAGCTGTAGACTTGTCCTAA 135
SMFIAACAGKKKTADKAVDLSK 480
GATGGTCTGCTAGGTGACATTCTACAGGATCTTAACACTGAGACACCTCAATAACTCC 155
DGLLGDILLQDLNTETPQITP 540
KCACCTGTAATGATACTGAAGAAGAAAGATCCATTGGAGCTTCACCGAATCCTTCTC 175
PPVMILLKKKRSLIGASPNPFS 600
GTGCACACCGCCAGGTCCTTCAGGAATAATTGCTTCCCTGTCTCCAGAAAGGA 195
VHTATAVVPSGKIAASPVSRE 660
CCTCCATTAACTCCTTCTTAAACGTGCTGAATTTGCTGGCGATGTACAGGT 215
PPLTPVPLKRAEFAGDDVQV 720
GAGAGTACAGAAGAGCAGGAGTCAGGGCAATGGAGTTTGAAGATGGTGACTTTGA 235
ESTEEEQESGAMEFEDGDFD 780
TAGCCCATGGAAGTTGAAGAGGTGGACCTGGAGCCCTATGGCTGCCAAGGCTTGGGACAA 255
EPMEEVEVDLEPMAAKAWDK 840
- 275

FIG. 3A

GAGAGTACAGAAGAGCAGGAGTCAGGGCAATGGAGTTTGAAGATGGTGACTTTGA

4/23

AGAGAGTGCCAGCAGGAAGTGAACAAGAGCGGATTCTGGAAAGGACCGTGTC 900
E S E P A E E V K Q E A D S G K G T V S - 295
CTACTAGGAAGTTTCTCCGGATGCTCTGTGGACATTGATCAAGAAGGTGATAG 960
Y L G S F L P D V S C W D I D Q E G D S - 315
CAGTTTCTCAGTCAAGAAGTTCAAGTGGATCCAGTCACCTCCCATTTGGTAAAGGGC 1020
S F S V Q E V Q V D S S H L P L V K G A - 335
AGATGAGGAACAAGTATCCACTTTTATGGTTGGATGCTTATGAGGATCAGTACAACA 1080
D E E Q V F H F Y W L D A Y E D Q Y N Q - 355
ACCAGGTGTGATTTCTGTTTGGAAAGTTTGGATTGAATCAGCCGAGACCCATGTGAG 1140
P G V V F L F G K V W I E S A E T H V S - 375
CTGTTGTGTCATGGTGAATAATCGAGCGACGCTTTACTTCTCCCGTGAATGAA 1200
C C V M V K N I E R T L Y F L P R E M K - 395
AATTGATCTAATACGGGGAAGAACAGGAACCTCCAATTTCAATGAAGGATGTTTATGA 1260
I D L N T G K E T G T P I S M K D V Y E - 415
GGAATTTGATGAGAAATAGCAACAATAAATAATGAAGTTCAAGTCTAAGCCAGT 1320
E F D E K I A T K Y K I M K F K S K P V - 435
GAAAGAACTATGCTTTTGAGATACCTGATGTTCCAGAAAATCTGAGTACTTGGAAGT 1380
E K N Y A F E I P D V P E K S E Y L E V - 455
TAAATACTCGGCTGAATGCCACAGCTTCCCTCAAGATTGAAAGGAGAAAACCTTTTCTCA 1440
K Y S A E M P Q L P Q D L K G E T F S H - 475
TGATTTGGGACCAACACATCTAGCCTGGAACCTGTTCTTGATGAACAGAAAGATCAAAGG 1500
V F G T N T S S L E L F L M N R K I K G - 495
ACCTTGTGGCTTGAAGTAAAGTCCACAGCTCTTGATCAGCCAGTCAGTTGGTGTA 1560
P C W L E V K S P Q L L N Q P V S W C K - 515
AGTTGAGGCAATGGCTTTGAAACAGACCTGGTGAATGTAATTAAGGATGTCAGTCCACC 1620
V E A M A L K P D L V N V I K D V S P P - 535
ACCGCTTGTCGTGCTTTCAGCATGAAGACAAATGCAGAAATGCAAGAACCATCAAAA 1680
P L V V M A F S M K T M Q N A K N H Q N - 555

FIG. 3B

SEQUENCE LISTING

5/23

PGAGATTATTGCTATGGCAGCTTTGGTCCATCACAGTTTTCATTGGATAAAGCAGCCCC 1740
 E I I A M A A L V H H S F A L D K A A P - 575
 AAAGCCTCCCTTTCAGTCACACTTCTGTGTGTCTAAACCAAGGACTGTATTTTCC 1800
 K P P F Q S H F C V V S K P K D C I F P - 595
 ATATGCTTCAAGAAGTCATTGAGAAAAAGATGTGAAGTTGAGGTTGCTGCAACAGA 1860
 Y A F K E V I E K K N V K V E V A A T E - 615
 AAGAACACTGCTAGGTTTTTTCCTTGCAAAAGTTCACAAAATTGATCCTGATATCATGT 1920
 R T L L G F F L A K V H K I D P D I I V - 635
 GGTCAATAATTTATGGGTTTGAAGTGAAGTACTACTGCAGAGAAATTAATGTGCAAA 1980
 G H N I Y G F E L E V L L Q R I N V C K - 655
 AGCTCCTCACTGGTCCAAGATAGGTGCGACTGAAGCGATCCAACATGCCAAAGCTTGGGG 2040
 A P H W S K I G R L K R S N M P K L G G - 675
 CCGGAGTGGATTGGTGAAAGAAATGCTACCTGTGGTCTGATGATCTGTGATGTGGAAT 2100
 R S G F G E R N A T C G R M I C D V E I - 695
 TTCAGCAAAAGGAATTGATTCTGTGTAAAGCTACCATCTGTCTGAACCTTGTTCAGCAGAT 2160
 S A K E L I R C K S Y H L S E L V Q Q I - 715
 ICTAAAAAAGGTTGTAATCCCAATGGAAAATATACAAAATATGTACAGTGAATC 2220
 L K T E R V V I P M E N I Q N M Y S E S - 735
 TTCTCAACTGTTATACCTGTTGGAACACACCTGGAAAGATGCCAAGTTCATTTTCAGAT 2280
 S Q L L Y L L E H T W K D A K F I L Q I - 755
 CATGTGTGAGCTAAATGTTCTTCATTAGCATTCAGATCACTAACAATCGCTGGGAACAT 2340
 M C E L N V L P L A L Q I T N I A G N I - 775
 TATGTCCAGGACCTGATGGGTGGAGATCCGAGCGTAACGAGTCTTGTGCTTCATGC 2400
 M S R T L M G G R S E R N E F L L L H A - 795
 ATTTTACGAAACAACATAATTGTGCTGACAAGCAGATTTTCAGAAAGCCTCAGCAAAA 2460
 F Y E N N Y I V P D K Q I F R K P Q Q K - 815
 ACTGGGAGATGAAGAAATTGATGGAGATACCAATAATAACAAGGACGTAA 2520
 L G D E D E E I D G D T N K Y K K G R K - 835

FIG. 3C

ADAPTIVE QUEST

6/23

SAAAGCAGCTTATGCTGGAGGCTTGGTTTGGACCCCAAGTTGGTTTATGATAAGTT 2580
K A Y A G G L V L D P K V G F Y D K F - 855
CATTGCTTCTGGACTTCAACAGTCTATATCCTTCCATCATTCAGGAATTTAACATTG 2640
I L L D F N S L Y P S I I Q E F N I C - 875
TTTTACAACAGTACAAGAGTTGCTTCAGAGGCACAGAAAGTTACAGAGGATGGAGAACA 2700
F T T V Q R V A S E A Q K V T E D G E Q - 895
AGAACAGATCCCTGAGTTGCCAGATCCAAGCTTAGAAATGGGCATTTTGCCCGAGAGAT 2760
E Q I P E L P D P S L E M G I L P R E I - 915
CCGGAACCTGGTAGAACGGAGAAAACAAGTCAAAACAGCTAATGAACAGCAAGACTTAA 2820
R K L V E R R K Q V K Q L M K Q Q D L N - 935
TCCAGACCTTATCTTCAGTATGACATTCGACAGAAGGCTTTGAAGCTCACAGCGAACAG 2880
P D L I L Q Y D I R Q K A L K L T A N S - 955
TATGTATGTTGCCCTGGGATTTCCCTATAGCAGATTTTACGCCAAACCACTGGCTGCCCT 2940
M Y G C L G F S Y S R F Y A K P L A A L - 975
GGTGACATACAAGGAAGGAGATTTTGATGCATACGAAGAGATGGTACAAAAGATGAA 3000
V T Y K G R E I L M H T K E M V Q K M N - 995
TCTTGAAGTTATTTATGGAGATACAGATTCAATTATGATAAACACCAATAGCACCAATCT 3060
L E V I Y G D T D S I M I N T N S T N L - 1015
SGAAGAAGTATTAAAGTTGGAAACAAGGTAAGTGAAGTGAATGAATGAATGTACAAACT 3120
E E V F K L G N K V K S E V N K L Y K L - 1035
GCTTGAAATAGACATTGATGGGGTTTCAAGTCTCTGCTACTGCTGAAAAAAGAAAGTA 3180
L E I D I D G V F K S L L L K K K Y - 1055
CGTGCCTGTTGTTGAGCCCAACGTCGGATGGGAATTATGTACCAACACAGGAGCTCAA 3240
A A L V V E P T S D G N Y V T K Q E L K - 1075
AGGATTAGATATAGTTAGAAGAGATTGGTGTGATCTTGCTAAGACACTGGAACTTTGT 3300
G L D I V R R D W C D L A K D T G N F V - 1095
GATTGCCAGATTCTTCTGATCAAGCCGGACACTATAGTGGAACATTCAGAAGAG 3360
I G Q I L S D Q S R D T I V E N I Q K R - 1115

FIG. 3D

7/23

GCTGATAGAAATTGGAGAAATGTGCTAAATGGCAGTGTCCAGTGAGCCAGTTTGAAT 3420
L I E I G E N V L N G S V P V S Q F E I -1135
TAACAAGGCATTGACAAAGGATCCCCAGGATTACCCTGATAAAAAAGCCTACCTCATGT 3480
N K A L T K D P Q D Y P D K K S L P H V -1155
ACATGTTGCCCTCTGGATAAAATCTCAAGGAGGCAGAAAGGTGAAGCTGGAGATACTGT 3540
H V A L W I N S Q G G R K V K A G D T V -1175
GTCATATGTCATCTGTCAGGATGGATCAAACTCACTGCAAGTCAGAGGGCCTATGCCGC 3600
S Y V I C Q D G S N L T A S Q R A Y A P -1195
TGAGCAGCTGCAGAAACAGGATAATCTAACCATTGACACCCAGTACTACTGCCCCAGCA 3660
E Q L Q K Q D N L T I D T Q Y Y L A Q Q -1215
GATCCACCCAGTCGTGGCTCGGATCTGTGAACCAATAGACGGAATTGATGCTGTCCTCAT 3720
I H P V V A R I C E P I D G I D A V L I -1235
TGCAACGTGTTGGGACTTGACCCCAACCAATTTAGAGTTCATCATTAATAAGATGA 3780
A T W L G L D P T Q F R V H Y H K D E -1255
AGAGAAATGATGCTCTACTTGGTGGCCAGCACAGCTCACTGATGAAGAGAAATACAGGA 3840
E N D A L L G G P A Q L T D E E K Y R D -1275
CTGTGAAGATTCAAAATGTCCATGCCCTACATGTGGAAGTGAAGTATTTATGATAATGT 3900
C E R F K C P C P T C G T E N I Y D N V -1295
CTTTGATGGTTCGGGAACAGATATGGAGCCCACTGTGATCGTTGCAGTAACATCGATTG 3960
F D G S G T D M E P S L Y R C S N I D C -1315
TAAGGCTTCACCTCTGACCTTTACAGTACAACTGAGCAACAAATTGATCATGGACATTAG 4020
K A S P L T F T V Q L S N K L I M D I R -1335
ACGTTTCATTAAAAGTACTATGATGGCTGGTTGATATGTGAAGAGCCAACTGTGCGCAA 4080
R F I K K Y Y D G W L I C E E P T C R N -1355
TCGAACTCGTCACCTTCCCTTCAATTCTCCGAACTGGCCCTCTTTGCCCGCCTGCAT 4140
R T R H L P L Q F S R T G P L C P A C M -1375
GAAAGCTACACTTCAACCAAGAGTATCTGACAAGTCCCTGTACACCCAGCTGTGCTTTA 4200
K A T L Q P E Y S D K S L Y T Q L C F Y -1395

FIG. 3E

SUBSTITUTE SHEET

8/23

CCGGTACATTTTGTGATGGGAGTGTGCACCTGGAGAACTTACTACCGATCATGAGAAAGA 4260
R Y I F D A E C A L E K L T T D H E K D -1415
TAAATGAAGAAGCAATTTTACCCCAAGTTCTGCAGGACTACAGAAACTCAAGAA 4320
K L K K Q F F T P K V L Q D Y R K L K N -1435
CACAGCAGAGCAATTCTTGTCGCCGAAGTGGCTACTCCGAAGTGAATCTGAGCAAACTCTT 4380
T A E Q F L S R S G Y S E V N L S K L F -1455
CGCTGTTGTGCCGTGAATCTTAAGGGAATCCCAGGAGTAACCAAGGAGGGGTAGTTG 4440
A G C A V K S * -1462
AAAAATCCCAGCTTCCTCTGTGCTCCACTCTGGCCCTAAATGCTCCTCCAGCATCTGTT 4500
TCTCCCTTGGGACTGTGTCTCATGTTTGTGTAATGTAGACCAGGAAGGGGCTGCAAA 4560
AATGTTGAGTCTAATGTTGTAAGCATCATAGAAATTCCTGTCTTCATATTAAGATGTAC 4620
TGCTTTAAACACAACTCCAGAGCCCTCCCAAGCTCCCCCTCCCAAGCTCCTGAAGAC 4680
CCGGTTCTGAGGGGAAATTGCTACTTGGATTGAGAGTAGCTGGAATGTAAGTGACC 4740
CCAGGCTTTGCTCAGGGCCTTTAGCCTATGTCCCCCCACATAAAGAGAGCTTCTCAGAG 4800
CCTGACTGAAGAGCTGACGTTTGTCTTTTTCATATGCCAATTAACCCGGTCTAAATCCA 4860
AATGCTTCTCCAGCCATCCAGGAGTGGCTGTCTTTTCAGTCTGTCTTTATATAGTA 4920
GCTGAGGGGAAGATTAGAAGCCTTGCACTCACTAAATAGATTAAACAGAGCAGGCTTG 4980
TTTGTGAAATTGCTCCAAAGTCCAAACAGACACACACTGAGCAGGTGTTTACACTCACAT 5040
TCCCTTTTGGCCCTTAAATAGAAAGTGCAAGGTAAGGTTTATACAAAGAAAGCACAT 5100
TGAAATAAATTGATACTCTAACAAATCCATTAAACATGTGTAGGGTTACGGTGAGGATCA 5160
TGTGTTGATTGCAAAAACGGGAGAGGGATGCTTAATTGGCCCTCGCTTGCTATTTTTT 5220
TCTCATTTCTTCACAAATAGGACCGTCTTTGGCAGCAGCAAAATGTATTTTCAGTATGGCAG 5280
TCTTCCCTCTTACATTATTGGTAAGATTATACTAACAAAATGTTTCCCTGTGACAAAT 5340
TATGCTGTGTTTTTAAAAAACATTGACCTGTGTGTTTTTATAAAGAAAGTATGTTGT 5400
GCCTTCTCTTAAGAAATAAGTTTCTAAAGGG 5433

FIG. 3F

வினாக்கள்: 1. பௌத்த

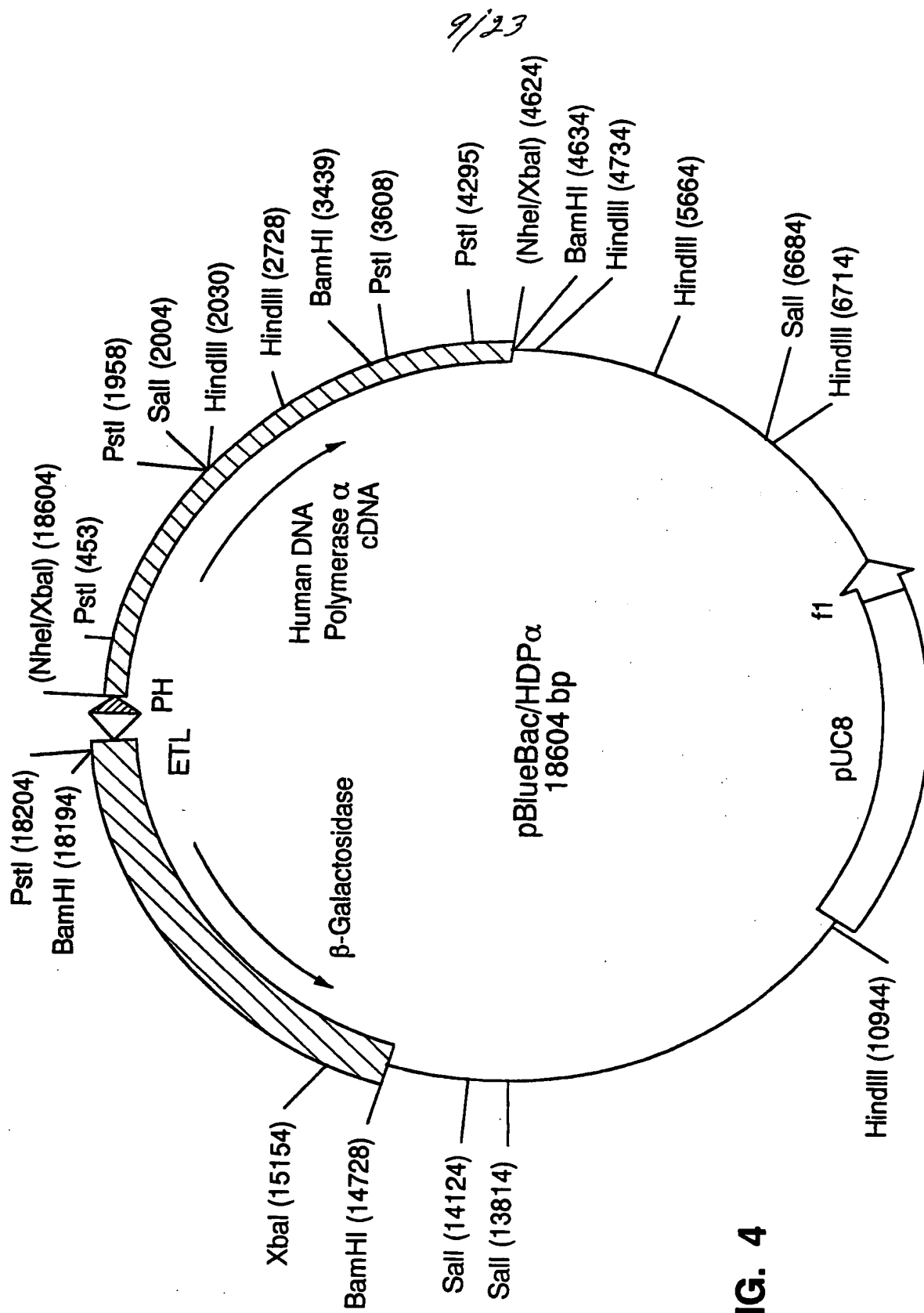


FIG. 4

10/83

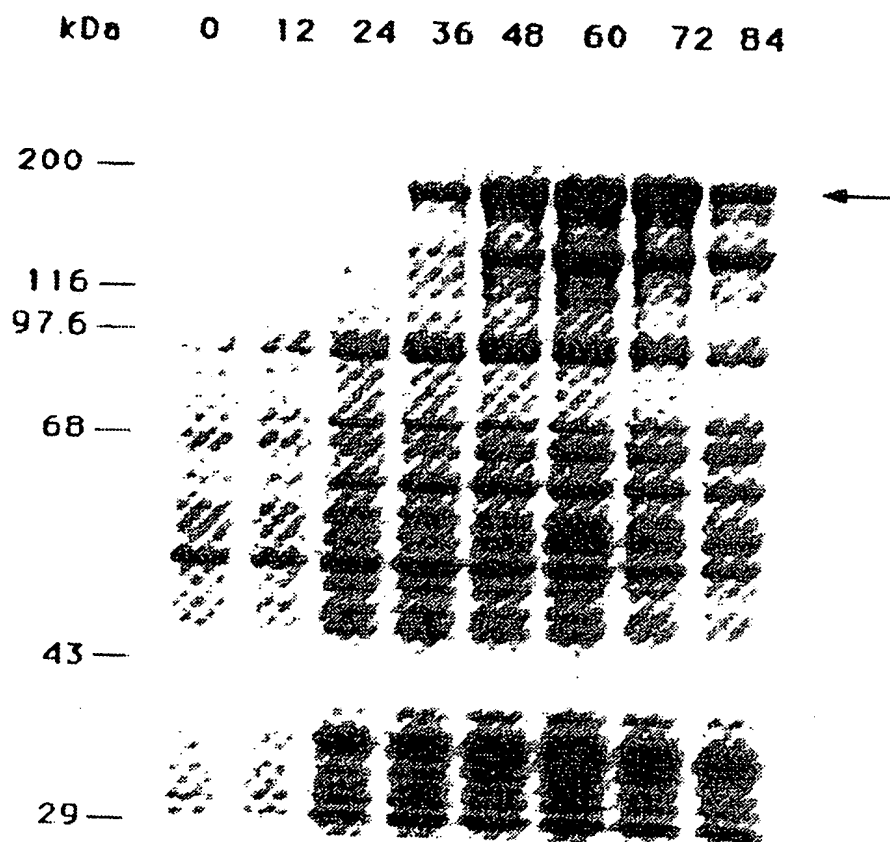


FIG. 5A

11/23

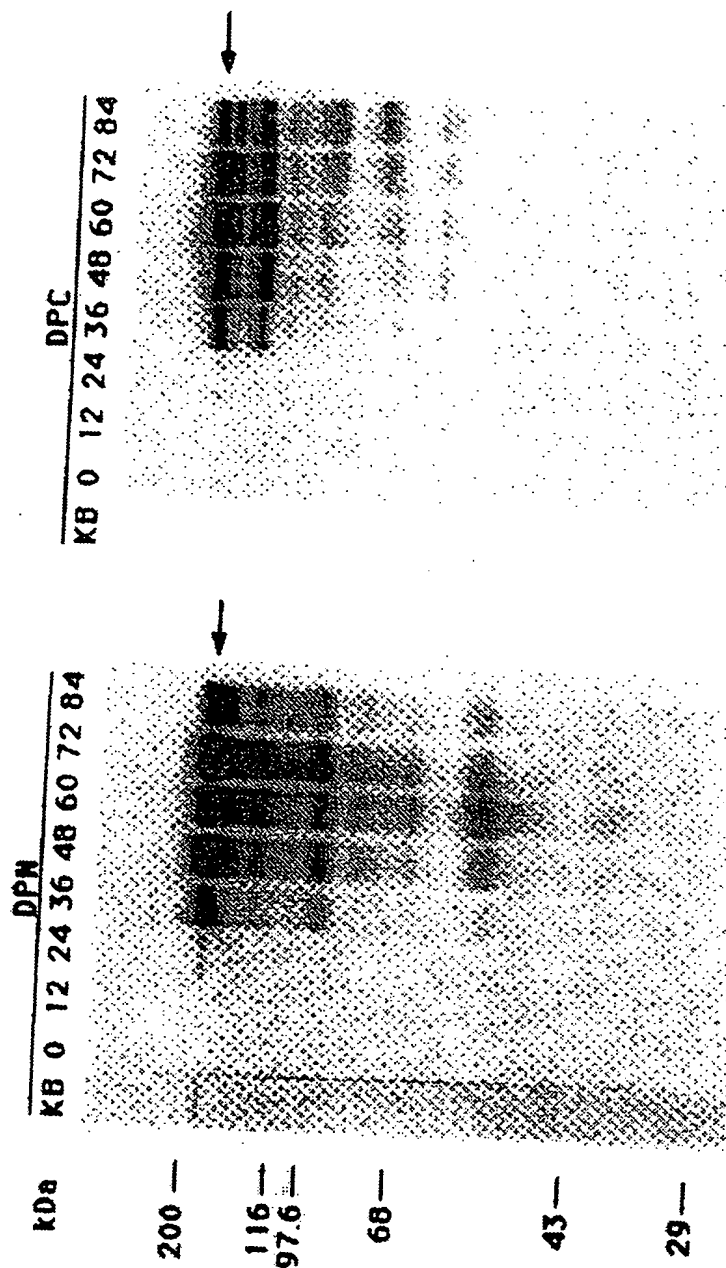


FIG. 5B

12/23

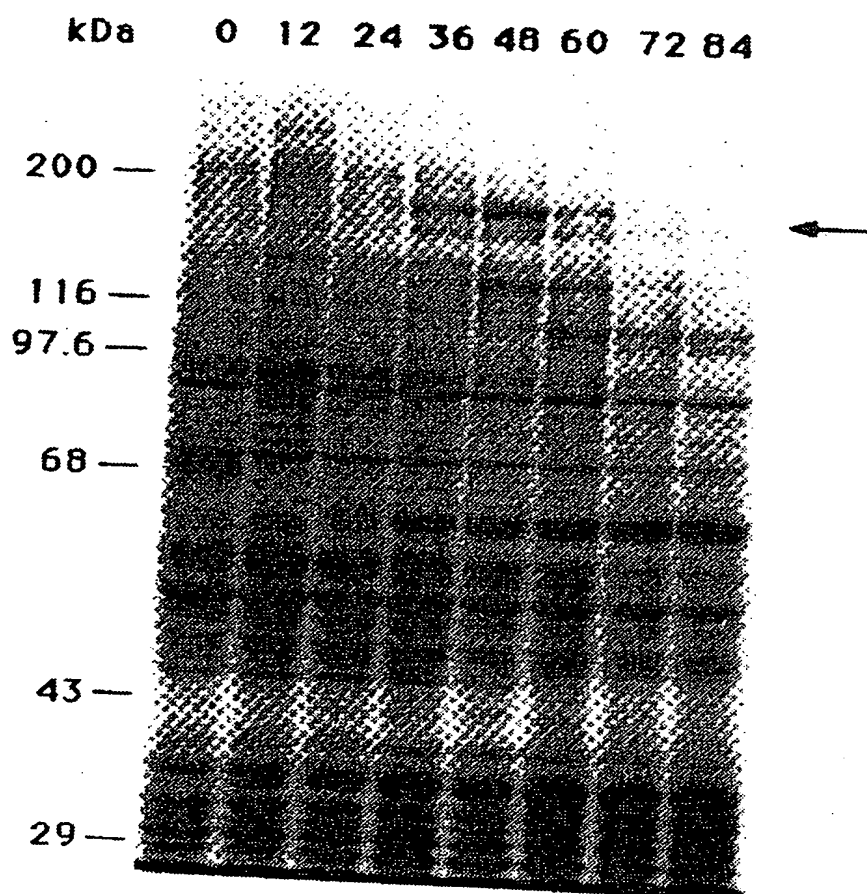


FIG. 5C

SUBSTITUTE SHEET

13/23

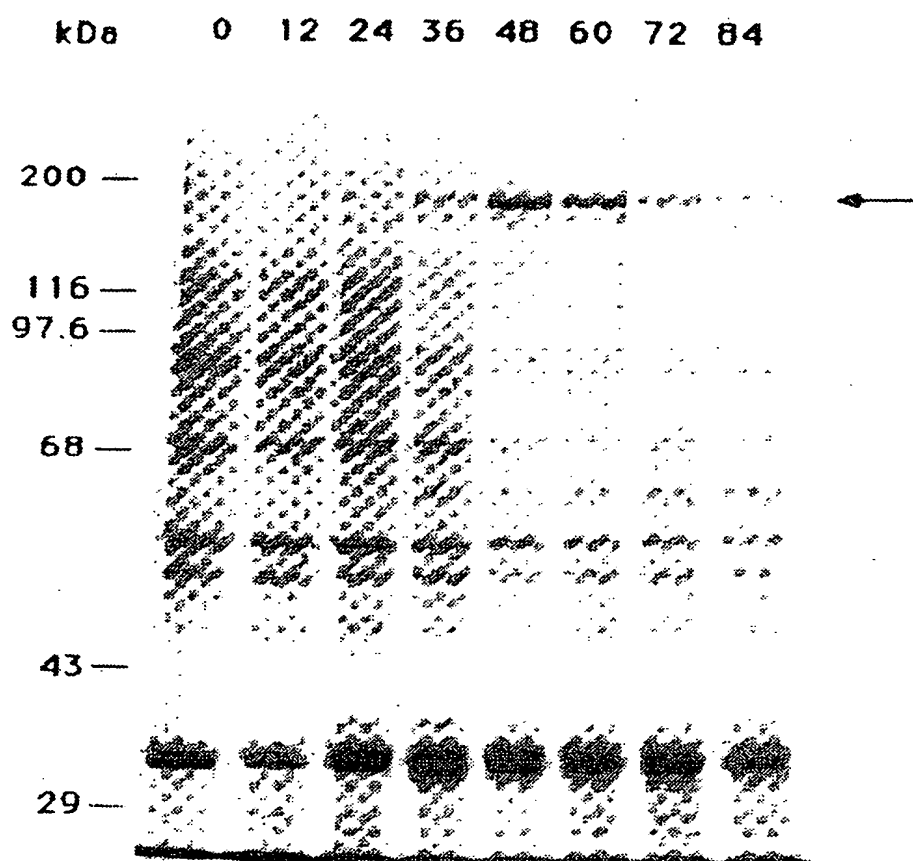


FIG. 5D

SUBSTITUTE SHEET

14/23

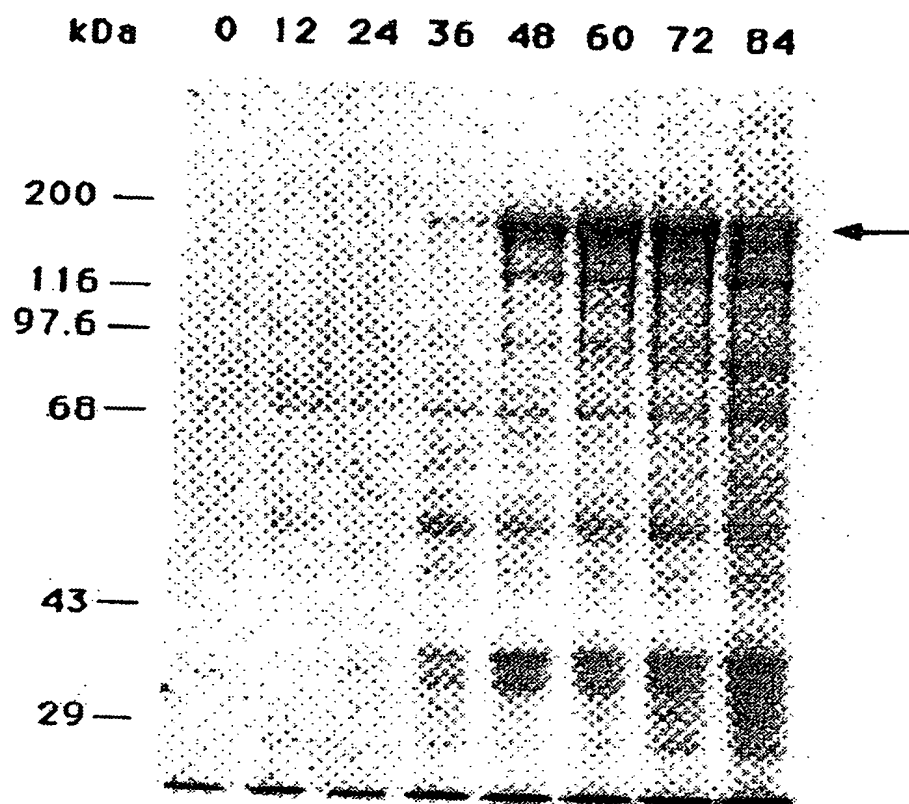


FIG. 5E

SUBSTITUTE SHEET

15/23

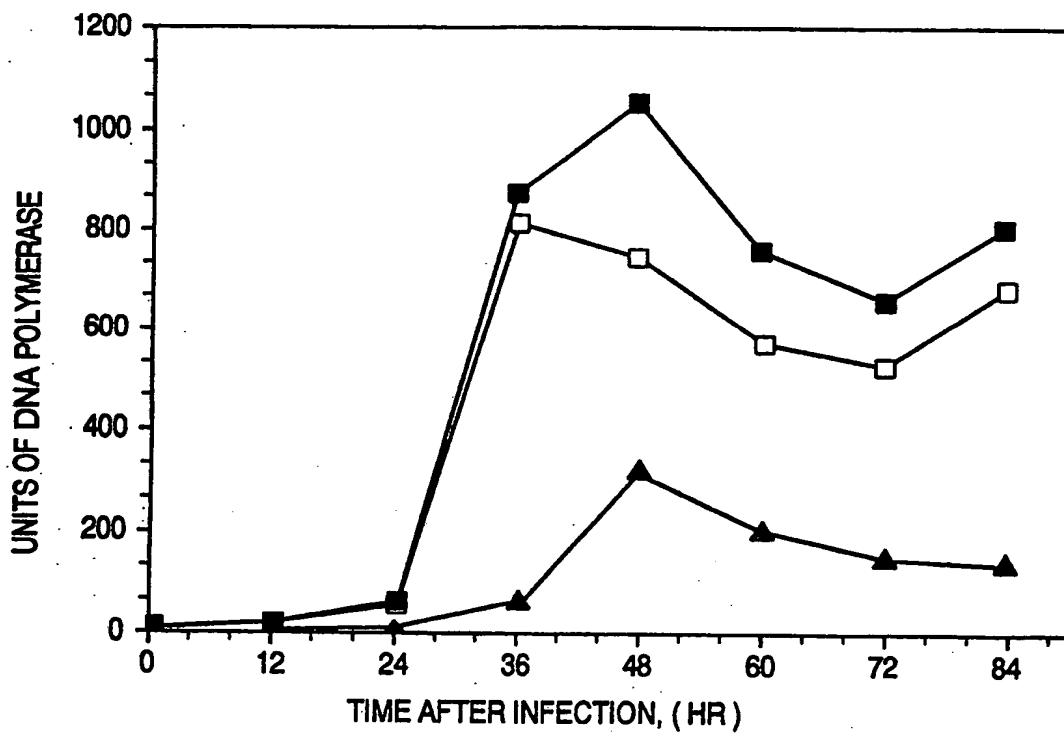
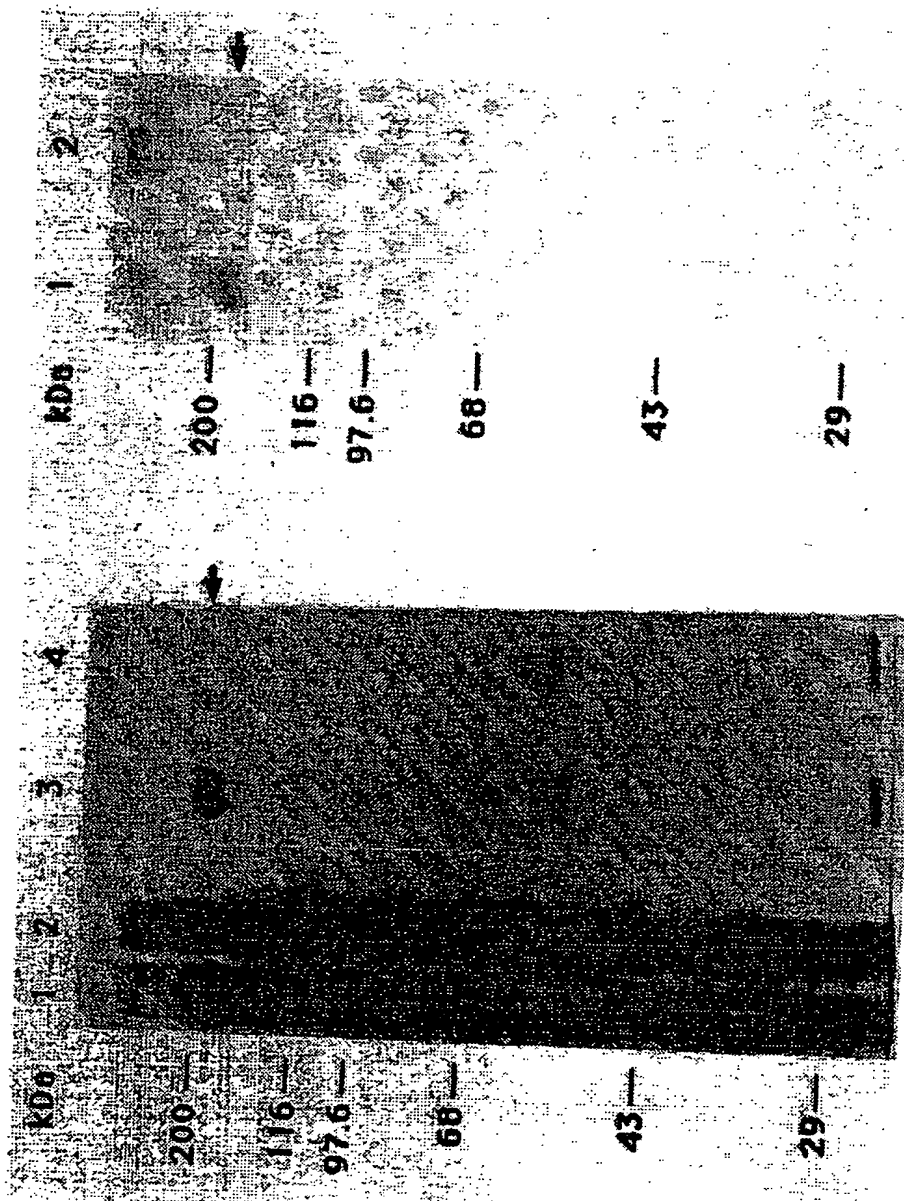


FIG. 5F

SUBSTITUTE SHEET

16/23



17/93

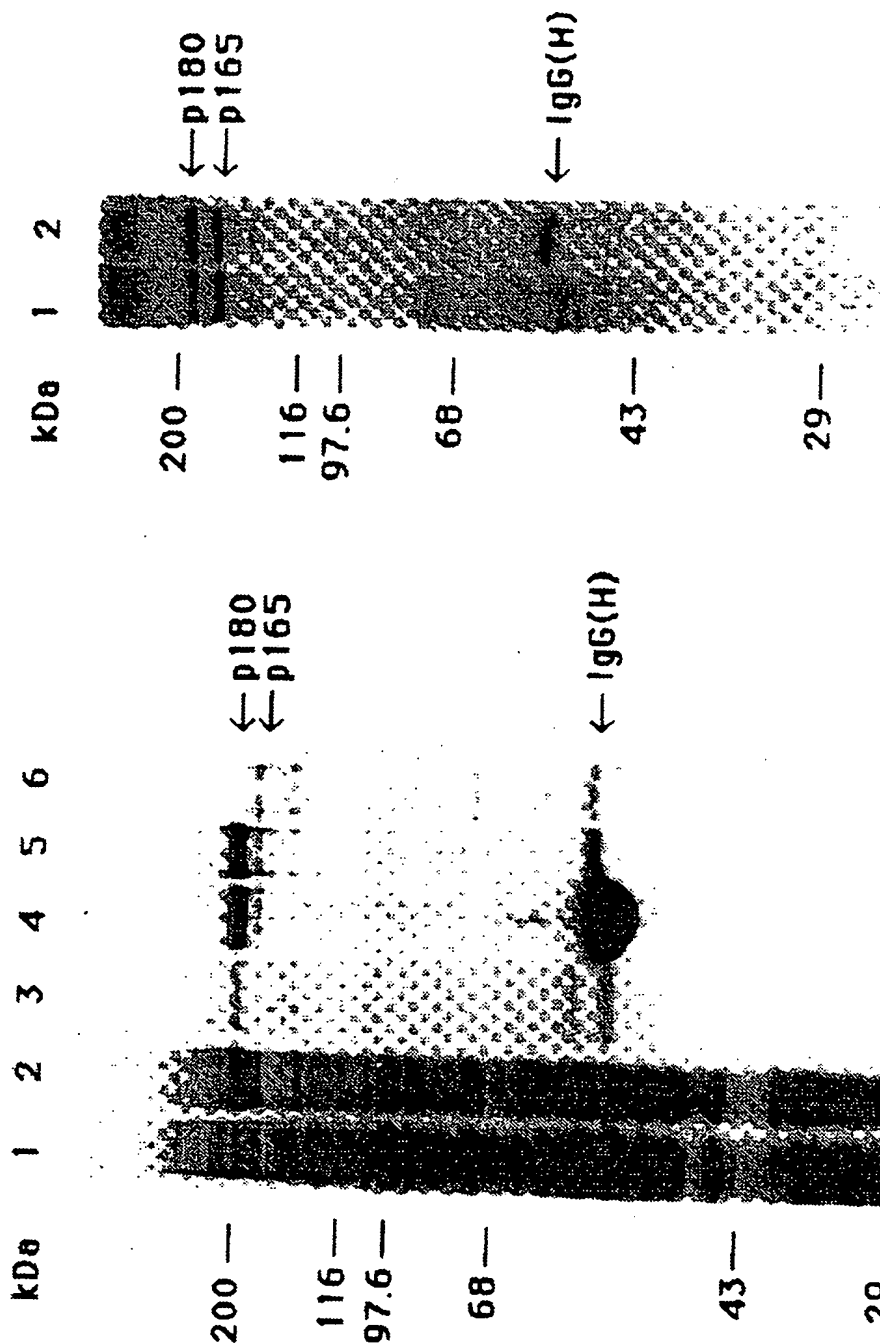


FIG. 7A

FIG. 7B

18/23

M13mp18 5'-AGGCATCGAAGCTTGGCACTGGCCGTCG-3'
TTCGAACCGTGACCGGCAGC-5'

TTTTTTT

1 2 3 4 5 6 7 8 9 10 11 12 13

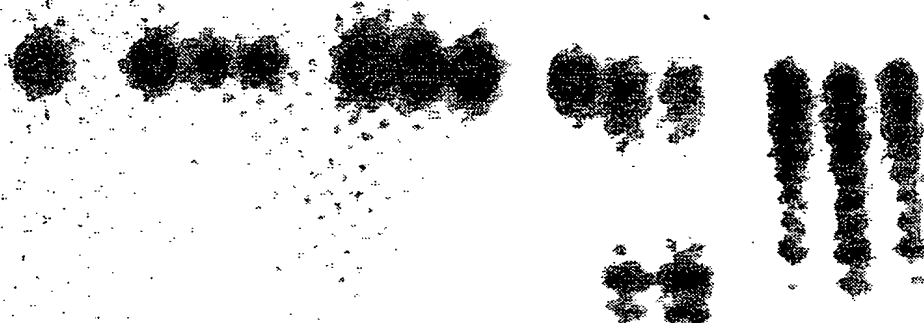


FIG. 8A

M13mp18 5'-GTCGTGACTGGGAAAACCCTGGCG-3'
CAGCACTGACCCTTTTGGGACCGC-5'

1 2 3 4 5 6 7 8 9 10 11 12 13

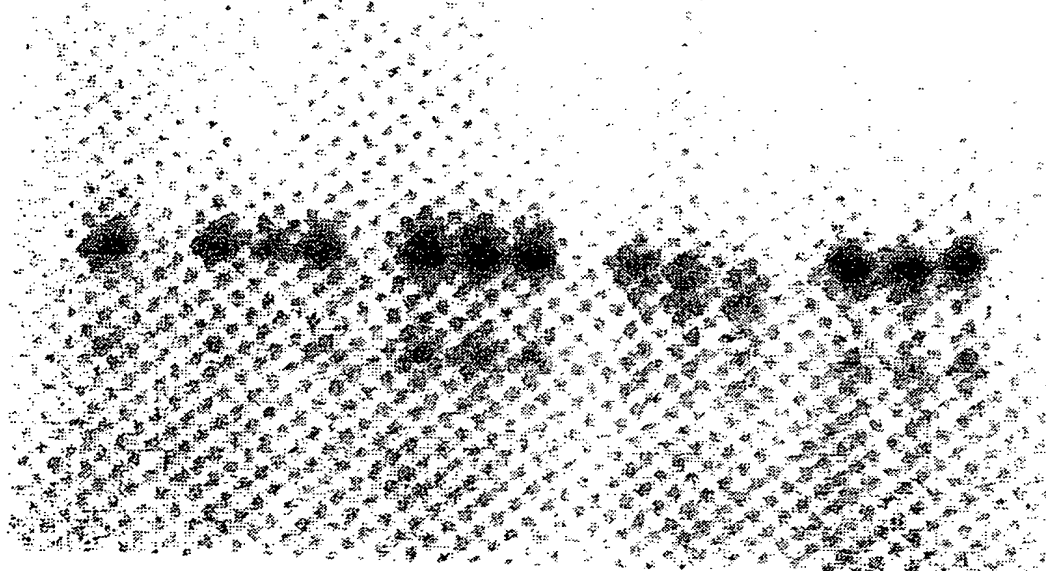


FIG. 8B

19/23



FIG. 9

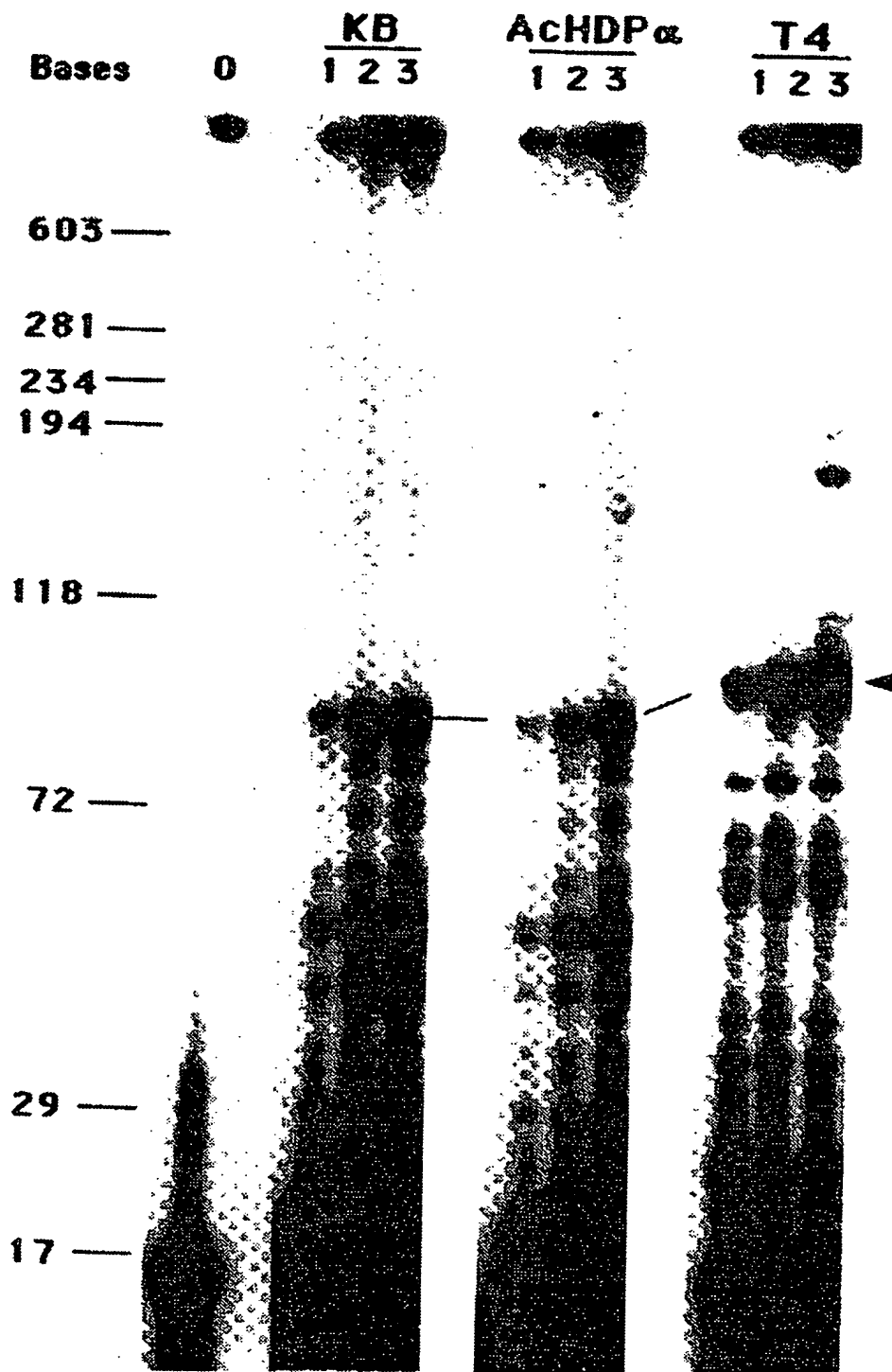
20/23

FIG. 10A
SUBSTITUTE SHEET

21/23

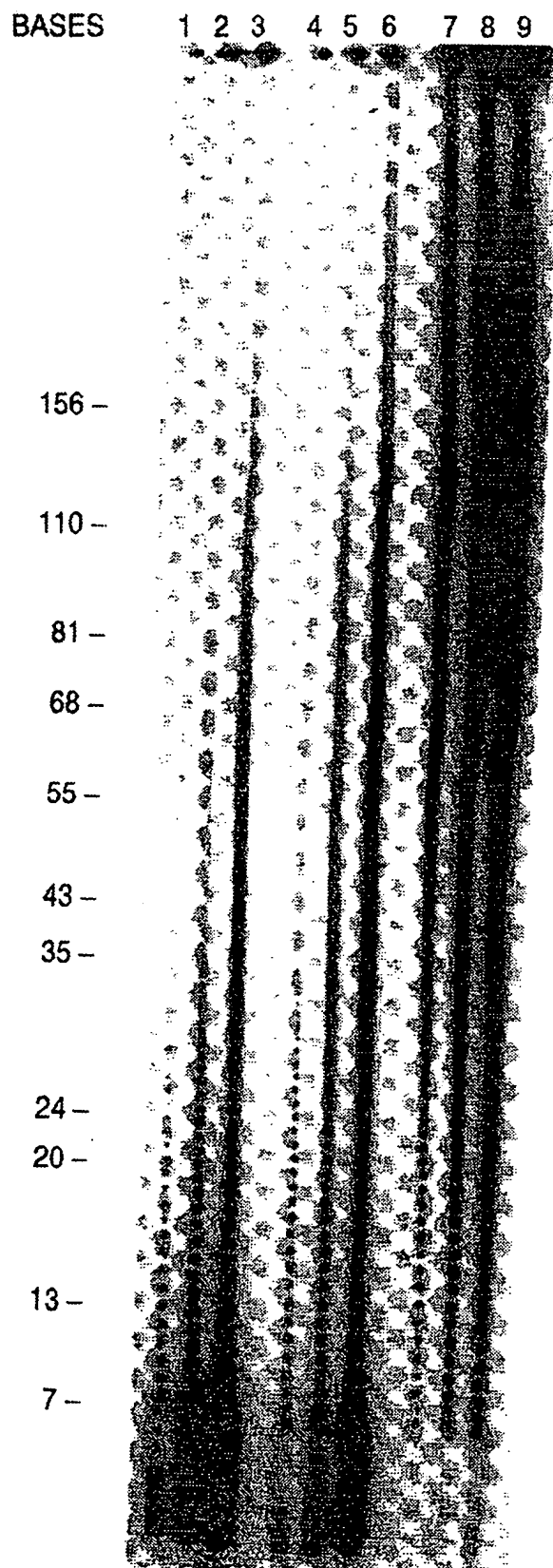


FIG. 10B

22/23

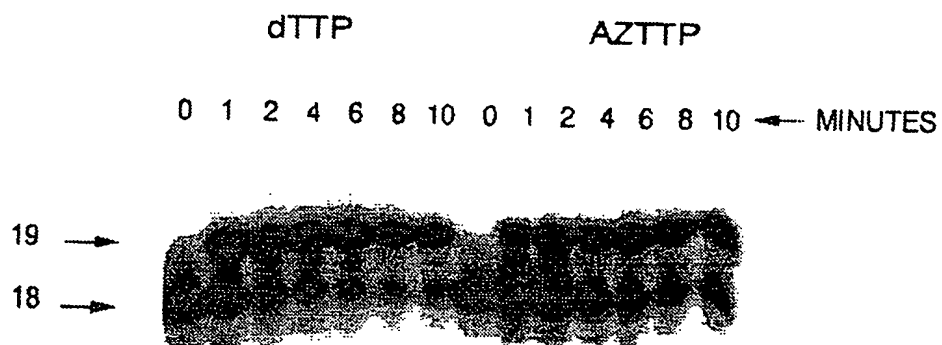


FIG. 11A

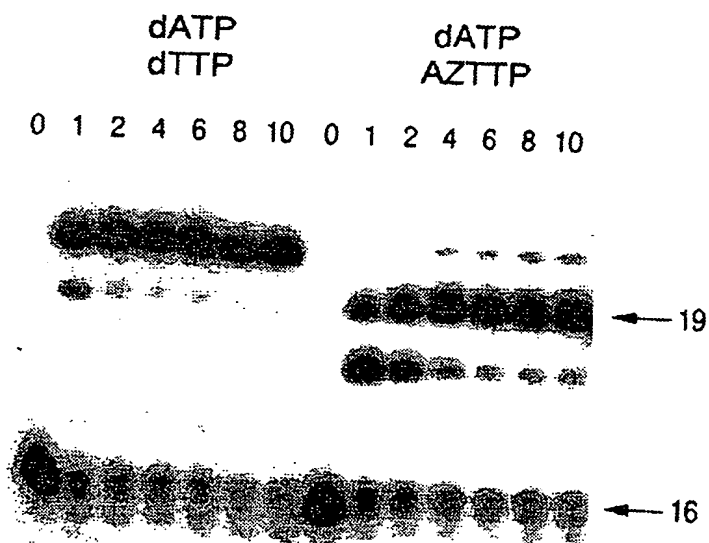


FIG. 11B

SUBSTITUTE SHEET

*23/23*Co-infection of SV40 Tag and human Pol α

Immunoprecipitated

Western

Coomassie stain

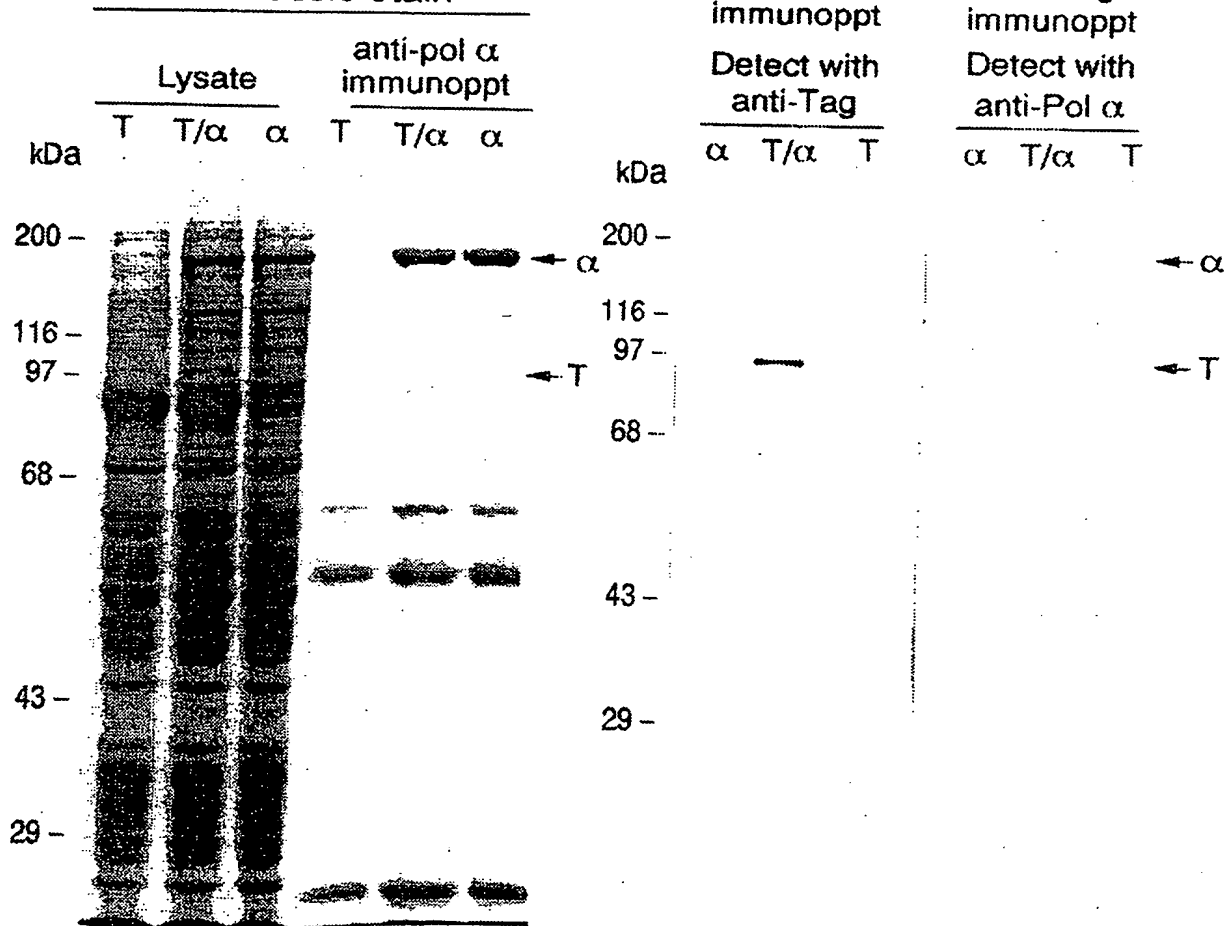


FIG. 12A

FIG. 12B

FIG. 12

SUBSTITUTE SHEET

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) : C07H 21/04; C12P 21/02; C12N 15/86; C12Q 1/48; C12Q 1/70

US CL : 536/27; 435/5, 15, 70.1, 320.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/27; 435/5, 15, 70.1, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
cas online, medline, APS, biosis, embase**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
x	EMBO Journal, volume 7, No. 1, issued 1988, S. W. Wong et al., "Human DNA Polymerase Alpha Gene Expression is Cell Proliferation Dependent and its Primary Structure is Similar to Both Prokaryotic and Eukaryotic replicative DNA Polymerases", pages 37-47, especially page 39.	1,2,6
y		3-5,7-14
y	US, A, 4,745,051 (Smith et al.) 17 May 1988, abstract and col. 18, line 41.	3-5,7-8
y	Biochim. et Biophys. Acta, Volume 699, issued 1982, M. Rochowska et al., "Inhibition of Calf Thymus DNA Polymerase Alpha and of Normal and Cancer Cell Growth by Butylanilinouracil and Butylphenylguanine", pages 67-73, especially 69.	9-14

☐ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be part of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Z" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 05 January 1993	Date of mailing of the international search report 26 JAN 1993
Name and mailing address of the ISA/ Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer KENNETH R. HORLICK Telephone No. (703) 308-0196